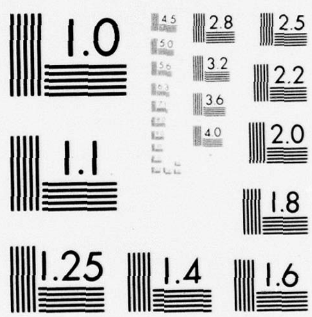
AD-A039 594 PENNSYLVANIA STATE UNIV UNIVERSITY PARK APPLIED RESE--ETC F/G 6/16
AN EXPERIMENTAL TREATMENT OF AUDITORY DISCRIMINATION OF COMPLEX--ETC(U)
NOV 76 C P JANOTA N00017-73-C-1418
UNCLASSIFIED TM-76-301 NL

1 OF 2
ADA039 594



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12)

AN EXPERIMENTAL TREATMENT OF AUDITORY

DISCRIMINATION OF COMPLEX NOISE-LIKE SOUNDS

Claus P. Janota

Technical Memorandum
File No. TM 76-301
November 19, 1976
Contract No. N00017-73-C-1418

Copy No. 11

The Pennsylvania State University
Institute for Science and Engineering
APPLIED RESEARCH LABORATORY
Post Office Box 30
State College, PA 16801

NAVY DEPARTMENT

NAVAL SEA SYSTEMS COMMAND

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TM-76-301	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER Technical memo.
4. TITLE (and Subtitle) AN EXPERIMENTAL TREATMENT OF AUDITORY DISCRIMINATION OF COMPLEX NOISE-LIKE SOUNDS.		5. TYPE OF REPORT & PERIOD COVERED PhD Thesis, June 1977
		6. PERFORMING ORG. REPORT NUMBER TM 76-301
7. AUTHOR(s) Claus Peter Janota 19 Nov 76		8. CONTRACT OR GRANT NUMBER(s) N00017-73-C-1418
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Pennsylvania State University Applied Research Laboratory P. O. Box 30, State College, PA 16801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Sea Systems Command Department of the Navy Washington, D. C. 20362		12. REPORT DATE November 19, 1976
		13. NUMBER OF PAGES 161 pages & figures
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 162p.		15. SECURITY CLASS. (of this report) Unclassified, Unlimited
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited, per NSSC (Naval Sea Systems Command) 1/3/77

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Whereas the factors important in the understanding of speech waveforms in the presence of extraneous noise or when distorted are reasonably well understood, factors important in recognizing noise-like sounds are not. Many of the noise-like sounds to which man is exposed serve to convey useful information and, under proper circumstances, allow the listener to make inferences about the source. The ability of the listener to react to differences between such sounds is significant in a number of industrial situations. This thesis deals

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

391 007

20. ABSTRACT (Continued)

with an attempt to quantify subjective responses to sounds, especially in the discrimination between similar sounding noise-like stimuli. A model of the discrimination process as one of feature extraction and decision making about the presence of salient features is proposed. The model and the signal patterns used in the experiments are restricted to features of the sounds which are dichotomous, i.e., present in one sound pattern and absent in the other of a pair of sounds to discriminate.

Experiments using both marine sounds and laboratory generated signals were conducted using trained University student listeners and Navy personnel. Two experimental methods were applied with one using a modification of classical threshold techniques. The other procedure used a continuous rating scale to elicit information about the confidence of the listener's discrimination decision.

The major portion of the data collected was for the case where the dichotomous feature was a band of noise. This features results either from filtering a marine sound or from deleting a band limited component from laboratory generated sound patterns. The data was analyzed from the point of view of hypothesis tests on the presence of the feature. Data available from signal detectability experiments was used to predict the required signal-to-noise ratio to make a discrimination decision. Good agreement was found between this prediction and experiment results.

The second dichotomous feature was that associated with quasi-periodic variations in the amplitude of the sound. The results for this feature also show good agreement with predictions. The data for two marine sounds and a third laboratory generated one exhibit consistent listener performance.

Results reported in the thesis cover a significant number of discrimination tasks using complex sounds. The findings support the model of the process as one of feature extraction and subsequent decision making based on the detectability of the features. A number of areas requiring further study were identified, and although some anomolies were noted, the work provides the basis for future detailed studies to further quantify the processes involved in the recognition of noise-like sounds by humans.

ACKNOWLEDGMENT

The author gratefully acknowledges the interest and encouragement of Dr. R. Farwell, his thesis advisor, and that of other members of the committee. The numerous suggestions by Drs. Martin and Michael are specifically noted.

In addition, the assistance rendered by the Applied Research Laboratory of The Pennsylvania State University under contract with the Naval Air Systems Command (AIR 370) is greatly appreciated.

The author also especially expresses his appreciation to the personnel at the Applied Research Laboratory, notably Mr. T. Way and Mr. R. Hoover, at the Naval Training Center, San Diego, and at the Naval Air Development Center, Warminster, who were instrumental in the preparation for, conduct of, and review of the experiments reported. Furthermore, the assistance of fellow graduate students F. McKendree and D. Martin in the data collection and reduction phases of the work was invaluable and warrants recognition.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

ABSTRACT

Whereas the factors important in the understanding of speech waveforms in the presence of extraneous noise or when distorted are reasonably well understood, factors important in recognizing noise-like sounds are not. Many of the noise-like sounds to which man is exposed serve to convey useful information and, under proper circumstances, allow the listener to make inferences about the source. The ability of the listener to react to differences between such sounds is significant in a number of industrial situations.

This thesis deals with an attempt to quantify subjective responses to sounds, especially in the discrimination between similar sounding noise-like stimuli. A model of the discrimination process as one of feature extraction and decision making about the presence of salient features is proposed. The model and the signal patterns used in the experiments are restricted to features of the sounds which are dichotomous, i.e., present in one sound pattern and absent in the other of a pair of sounds to discriminate.

Experiments using both marine sounds and laboratory generated signals were conducted using trained University student listeners and Navy personnel. Two experimental methods were applied with one using a modification of classical threshold techniques. The other procedure used a continuous rating scale to elicit information about the confidence of the listener's discrimination decision.

The major portion of the data collected was for the case where the dichotomous feature was a band of noise. This feature results either from filtering a marine sound or from deleting a band limited component from laboratory-generated sound patterns. The data was analyzed from the point of view of hypothesis tests on the presence of the feature. Data available from signal detectability experiments were used to predict the required signal-to-noise ratio to make a discrimination decision. Good agreement was found between this prediction and experiment results.

The second dichotomous feature was that associated with quasi-periodic variations in the amplitude of the sound. The results for this feature also show good agreement with predictions. The data for two marine sounds and a third laboratory-generated one exhibit consistent listener performance.

Results reported in the thesis cover a significant number of discrimination tasks using complex sounds. The findings support the model of the process as one of feature extraction and subsequent decision making based on the detectability of the features. A number of areas requiring further study were identified, and although some anomalies were noted, the work provides the basis for future detailed studies to further quantify the processes involved in the recognition of noise-like sounds by humans.

CHAPTER I

INTRODUCTION

1.1 Statement of the Problem

Many of the noise-like sounds to which man is exposed serve also to convey useful information. Under proper conditions, such noises allow the listener to identify the source and to make inferences about it. The ability to react to differences between such sounds is of considerable significance in a number of situations ranging from the estimation of a vehicle's speed to the adjustment of the cutting rate of a lathe. This thesis deals with an attempt to quantify subjective responses to sounds, especially in the discrimination between similar sounding noise-like stimuli. Relevant knowledge is reviewed and signal detectability is used as the basis for a model of the discrimination process. An experimental method is described and the results of studies using University students and Navy sonar operators are presented.

1.2 Importance of the Problem

The ability of a listener to recognize noise-like sounds, to respond to changes of such sounds, and to discriminate between complex sounds plays a role in man's everyday life (Sheridan and Ferrell, 1974). In industrial environments, this capability is vital to the safety and often the productivity of the worker

(Murrell, 1965). Efforts to protect workers from excessive industrial noise have resulted in mandated use of hearing protectors in many industrial situations (EPA, 1972). The effect of these protectors on the ability of the worker to respond to industrial noises can be determined if a thorough knowledge of the auditory process is available. In addition, future monitoring of industrial processes may rely on automated systems for sonic processing to remove humans from areas with excessive noise levels. Such automatic recognition and discrimination devices may well use concepts derived from speech recognition and understanding systems (Lesser et al., 1975; Myasnikov and Myasnikova, 1970). In order to make this transfer, a quantitative knowledge of the factors which allow humans to distinguish between noise-like sounds will be needed.

A large body of literature is concerned with the processes involved in speech recognition and discrimination (Licklider et al., 1948; Pollack, 1948; Stevens and House, 1972). However, the important aspects of noise-like sound recognition have not been investigated systematically.

1.3 Approach

This thesis describes an experimental approach designed to elicit listener decisions about which of two trained exposure sound patterns is subsequently presented as a probe in a noise background. The probe is slowly brought out of the noise until the listener has sufficient information to make a decision. The signals used were either marine sounds or laboratory-generated sounds which have

features in common with the natural ones. The two sounds differed in the presence or absence of a recognizable component or feature. The experiment, therefore, deals with the performance of listeners in distinguishing between two sounds which differ by virtue of a dichotomous feature.

A model of the discrimination performance in terms of a simple hypothesis test about the presence of the salient feature is presented in Chapter II. There, the case of a dichotomous feature which is a band-limited noise is treated in greater detail making use of signal detection concepts and Bayesian hypothesis testing. The concept is then expanded in Chapter III to the decision process involved in sequential hypothesis tests where the listener has the option to defer the terminal decision. It is seen that the criterion the listener uses is a very significant factor in the decision.

The design of the experiment is discussed in Chapter IV. Pre-recorded test sequences were used, and the subjects used either an A/B response recorder to signal a terminal decision or a continuous rating scale to report their degree of confidence and tentative discrimination decision as a function of time. Then, in Chapter V, the rationale for choosing certain sound patterns is explained. The decision to use marine sounds was based primarily on the fact that such sounds are of known significance in auditory discrimination, have many characteristics in common with sounds occurring in various industrial settings, and were expected to be familiar to the Navy personnel who were to serve as subjects in some of these experiments. The dichotomous features explicitly

addressed were bands of noise and temporal shifts in amplitude. Chapter V also describes the training of subjects and controls exercised to insure listener safety.

The discrimination of sound patterns with bands of noise dichotomous was one of the principal topics of investigation. Experimental data collected by others about the detectability of noise bands provides the basis for comparing a model of the discrimination task as one of feature extraction and detection to the observed results. The data is presented in Chapter VI along with analyses of the results from the point of view of signal detectability. The results show that the simple model is adequate for predicting performance in some of the experiments using complex marine sounds. The model must be extended, however, whenever the experiment is designed to elicit discrimination decisions with the probe signal pattern for which the dichotomous feature is absent. For this situation, other auditory clues must be used in the decision process.

Experimental results for sound patterns which differ in their temporal characteristics are presented in Chapter VII. The dichotomous feature (Ω_{AM}) is in that case the time varying amplitude of the waveform. This dichotomous feature also resulted in listener discrimination performance which can as a first approximation be explained as a feature detection.

The experimental technique used does introduce some procedural biases which must be taken into account whenever the results are analyzed. These biases and anomalies are discussed in Chapter VIII. One of the most difficult interactions to take into account in

experiments as this one is that of time on the memory of the listener and also on his criterion. There is strong evidence to indicate that even trained Navy sonar operators find it difficult to maintain a constant decision criterion in tests where the signal-to-noise ratio is slowly increased with time. The results of continuous rating experiments is presented in Chapter VIII as a means for mapping out the receiver operating characteristic (ROC) in a sequential discrimination task. The effects of detectability of the dichotomous feature and listener confidence are seen to yield results consistent with expectations.

CHAPTER II

THE AUDITORY RECOGNITION TASK

2.1 General

In this chapter, the problem posed by the need of a listener to recognize an auditory stimulus is investigated. Such recognition implies that the listener has first somehow been convinced of the presence of a sound source and must now decide which of a number of possible classes of sounds the source represents. In practice, the processes of detection and classification are closely related. This relationship will be seen to be especially intimate when the classification reduces to distinguishing between two sounds which differ in certain precisely defined ways. Also, in most cases of practical interest, the sound will be heard in the presence of extraneous noises which will tend to mask the presence of the sound of interest as well as to confound the recognition or classification task.

Figure 1 shows the essential elements of the recognition task in a concise way using a general estimation model (Van Trees, 1968) modified to include some concepts from pattern recognition (Fu, 1968). The model consists of the following four components:

a. Source Characterization. An acoustic source is characterized by a set of parameters. These parameters $\bar{a}$ define a

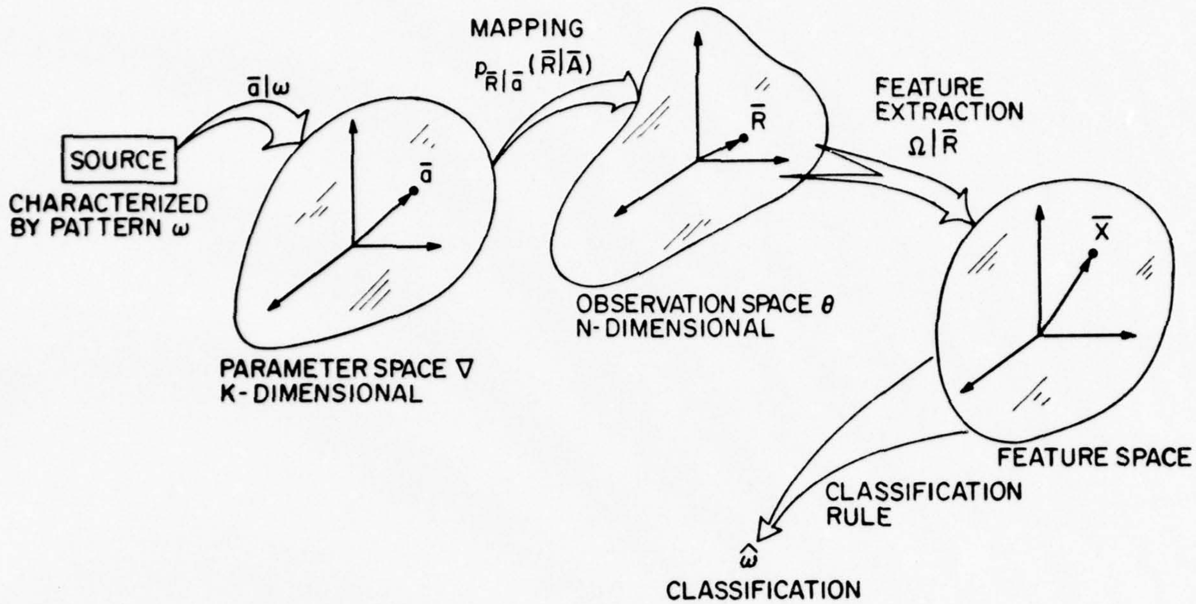


Figure 1. General Model of Classification Decisions with Auditory Cues.

multi-dimensional parameter space ∇ and they may be associated with such features of the sound as the presence of a band of noise, amplitude modulation, or temporal shifts in the sound's spectrum. Denote the acoustic pattern corresponding to acoustic source j as ω_j and the features as Ω_{ij}.

b. Observation. The source parameters map into an observation space θ according to some probabilistic law. The observable $\bar{R}$ is the listener's response to the stimulus. This observation includes the effect of the auditory transducer as well as subsequent neural and higher level cognitive processes.

c. Feature Extraction. It is hypothesized that the listener perceives the set of observables as features (Zhukov and Christovitch, 1974). These subjective features may not be equivalent to features Ω_{ij} above, however, and can be the result of physiology or cognition influenced by listener experience and training (Plomp and Smoorenburg, 1970; Reed, 1973; Ahumada et al., 1975). An N dimensional observation space reduces to an I dimensional feature space.

d. Classification Rule. It is further hypothesized that the listener will establish a measure of similarity between the set of observed features comprising a perceived pattern with patterns previously learned. On the basis of this measure, an identification of the sound results by applying

classification rules. Denote the classification resulting from the observables $\hat{\omega}(\bar{R})$.

In most listener classification tasks, the sound patterns ω_j and hence parameters $\bar{a}$ occur with some a priori probability. The probability of occurrence of various sound patterns strongly influences the classification problem by restricting the number of patterns which the listener expects to hear and by effecting the classification rule.

2.2 Patterns with Dichotomous Features

As a subset of the general classification model, consider the case of two signal patterns only. The task then reduces to one of deciding between two previously learned patterns on the basis of the observation. Such a decision may result when a lathe operator must decide between the sounds of a properly adjusted machine and one for which the cutting rate is excessive. In more complex classification tasks, it may be practical to arrive at a final classification as the result of two-way choices defining a decision tree (Reed, 1973).

Whenever the two sounds differ by the absence of specific features in one case where they occur in the other, the features are dichotomous. If only a single dichotomous feature Ω_i distinguishes two sound patterns ω_1, ω_2, the listener is faced with the following choices:

a. When the observation is thought to be associated with the presence of feature Ω_i, say that signal pattern ω_1 is heard.

b. Else conclude that signal pattern ω_2 is heard.

Stating the above decisions in terms of statistical hypothesis tests, define the two hypotheses:

H_0: Feature Ω_i is absent, decide pattern ω_2

H_1: Feature Ω_i is present, decide pattern ω_1.

In the limit where sound pattern ω_1 consists of a single feature Ω_1, the discrimination problem reduces to that of detection of the feature. While the detection problem is certainly of interest, this thesis is concerned with the classification or discrimination task instead. The detection problem has been treated elsewhere (Green, 1960a; Gundy, 1961; Stallard and Leslie, 1974; Swets, 1964), but it is instructive to state some of the applicable concepts here.

It is reasonable to suspect that the acceptance of hypothesis H_1 above requires at least an opportunity for the listener to detect the presence of the dichotomous feature Ω_i. It will be seen in Chapters VI and VII that, in some experimental cases, the detection of a dichotomous feature appears to be an adequate basis for a classification decision.

2.3 The Theory of Signal Detectability

The theory of signal detectability is a mathematical model of the mapping of a set of stimuli, consisting either of noise alone or from signal plus noise, into a two-point decision space: "no" the signal was not contained in the input, or "yes" the signal was contained in the input (Tanner and Sorkin, 1972). While these models are not intended to be descriptions of the way a human

observer makes this binary decision, they are normative models with which the observer's performance can be compared.

The models comprising this theory are applicable to any receiver operating on a set of inputs and as such, do not take into account the properties of the human per se. A number of workers have successfully demonstrated the utility of the theory of signal detectability of visual as well as auditory stimuli (Green and Swets, 1966; Peterson, Birdsall and Fox, 1954; Swets, 1964). It is generally found that the human performance falls short of that predicted for an ideal detector.

When applied to detection of marine sounds, the performance with broadband noise applies. This is due to the fact that marine sources radiate a broadband spectrum with perhaps some significant tonals (Kinsler and Frey, 1962). Industrial noise sources have similar characteristics (Ward and Fricke, 1969). The problem of detecting the presence of various broadband components reduces to a test of hypotheses about signals which are only known statistically. That is, the exact phases of the components, their instantaneous amplitudes, bandwidths, etc., are at best known in terms of their averaged power spectra. The theory of signal detectability for broadband features leads to a likelihood ratio test for signals known statistically. If we further assume that the noise-like sound can be modeled as a collection of samples with identical Gaussian probability densities, the ideal processor is an energy detector (Van Trees, 1968).

In accordance with the sampling theorem in the time domain, a band-limited signal of bandwidth W existing over a time interval T can be exactly determined by $2WT$ samples (Shannon and Weaver, 1949). These samples can be chosen to be statistically independent if the process is Gaussian (Lathi, 1968). The likelihood ratio test is then given by

$$\chi(\overline{R}) = \sum_{i=1}^{2WT} R_i^2 \underset{H_0}{\overset{H_1}{\gtrless}} \gamma = L(R^2) \quad , \tag{2.1}$$

where the threshold γ includes all sample invariant factors as well as the criteria (Van Trees, 1968). It is seen that the likelihood ratio test consists of computing the sum of squares of the statistically independent sample data points and comparing this to an appropriate threshold. Since the R_i's are Gaussian random variables, the sufficient statistic $\chi(\overline{R})$ is a random variable with a gamma distribution

$$p_\chi(\chi) = \frac{1}{\Gamma(d)\beta^d} \chi^{d-1} e^{-\chi/\beta} \quad , 0 < \chi < \infty \quad , \tag{2.2}$$

The parameters of the gamma distribution correspond to:

$d = \nu/2 = WT$, $\beta = 2\,\sigma_h^2$ (Hogg and Craig, 1970).

The parameter σ_h^2 is given by:

$\sigma_h^2 = \sigma_n^2$ under H_0 if R is $n(0, \sigma_n^2)$

$\sigma_h^2 = \sigma_n^2 + \sigma_s^2$ under H_1 where R is $n(0, \sigma_n^2 + \sigma_s^2)$.

The notation $n(\mu, \sigma_s^2)$ is used to denote a random variable with Gaussian probability density which has a mean μ and a variance σ^2.

Equation (2.2) reduces to that of a chi-square density function with ν degrees of freedom whenever $\sigma_h^2 = 1$.

Furthermore, whenever the number of degrees of freedom

$$\nu = 2WT > 100 \ ,$$

the random variable $\chi(R^2)$ can be closely approximated by a Gaussian density with the following parameters under H_0 and H_1 (Abramowitz and Stegun, 1964):

$$p_{\chi|H_0}(X|H_0) \text{ is } n(2WT\ \sigma_n^2,\ 4WT\ \sigma_n^4) \tag{2.3}$$

$$p_{\chi|H_1}(X|H_1) \text{ is } n[2WT\ (\sigma_n^2 + \sigma_s^2),\ 4WT\ (\sigma_n^2 + \sigma_s^2)^2] \ . \tag{2.4}$$

The performance of the hypothesis test can be completely characterized by the quantity d' (Van Trees, 1968). The detectability d' is defined as the distance between means on the two hypotheses when the variance has been normalized to unity (Swets, 1964). An identical definition is

$$(d')^2 = \frac{[E(X|H_1) - E(X|H_0)]^2}{1/2\ [Var(X|H_1) + Var(X|H_0)]} \ , \tag{2.5}$$

where the variances under the two hypotheses are different.

Substituting the mean values $E(X|H_1)$, $E(X|H_0)$ and variances $Var(X|H_1)$, $Var(X|H_0)$ for the Gaussian signal case, obtain:

$$d' = \left\{ \frac{4(WT)^2\,\sigma_s^4}{4WT\,\sigma_n^4 + 4WT\,\sigma_n^2\,\sigma_s^2 + 2WT\,\sigma_s^4} \right\}^{1/2}$$

$$d'_{opt} = (WT)^{1/2} \frac{\sigma_s^2}{\sigma_n^2[1/2(\sigma_s^2/\sigma_n^2)^2 + (\sigma_s^2/\sigma_n^2) + 1]^{1/2}} \tag{2.6}$$

where the notation d'_{opt} has been introduced to indicate theoretically optimal performance in the hypothesis test. The human observer will, in general, not attain this level of performance (Green and Swets, 1966).

For the small signal-to-noise case where

$$\sigma_s^2/\sigma_n^2 \ll 1 \quad ,$$

Equation (2.6) reduces to,

$$d'_{opt} = (WT)^{1/2}\,\sigma_s^2/\sigma_n^2, \quad \sigma_s^2/\sigma_n^2 \ll 1 \quad . \tag{2.7}$$

2.4 Discrimination of Complex Patterns

The above results for the detectability of a band of noise with bandwidth W is now extended to the discrimination task. Whenever two sound patterns differ by a single feature Ω_i, and this feature is a band-limited noise, feature Ω_i has an associated detectability d'_{opt} given by Equation (2.6). An equivalent measure of detectability can in theory be written for all features which distinguish the two sound patterns.

Although the distinguishing features in any real discrimination task may not be dichotomous, the sound patterns will differ in the relative detectabilities of the features. Also, whenever a number of

features comprise a sound pattern, it can be expected that some features will occur only when others are present. The features are, therefore, correlated one to another to some extent.

The general discrimination or recognition task for everyday sound patterns is often one involving multiple interacting features with a range of detectabilities. The analysis of the general task is extremely complex for this reason. This thesis deals, therefore, with a manageable subset of such sound patterns.

Even in the simple case of a few dichotomous features with controlled interactions, it can at the outset only be assumed that the detectability of the features will in some way be related to the discrimination performance.

The experimental technique discussed in later chapters was designed to test how accurately a model of the discrimination performance based on statistical hypothesis tests on dichotomous features agrees with human performance. Equation (2.1) for the likelihood test shows that the acceptance of hypothesis H_1 depends on a threshold γ as well as on the detectability of the signal. The performance of the ideal discriminator, and presumably that of the suboptimal human observer, can be changed by readjusting the threshold. The threshold is determined by the decision rules and takes into account the observer's knowledge about the relative frequency of occurrence of the signal pattern with the feature present (Egan, 1975).

CHAPTER III

SEQUENTIAL RECOGNITION

3.1 Deferred Decisions in Recognition

In many recognition tasks of practical interest, the confounding effect of extraneous noise precludes an immediate classification decision. For discrimination between signals with dichotomous features, this deferral of a decision is equivalent to sequential hypothesis testing (Wald, 1947).

At any time t_j, the listener is confronted with three possible decisions:

a. On the basis of the aural stimulus, decide H_0, the feature Ω_i is absent.

b. Given the aural stimulus, decide H_1, the feature Ω_i is present.

c. Defer the classification decision because there is insufficient evidence to decide either H_0 or H_1.

The listener will choose his course of action depending both on the observables and on his mental set (Leeper, 1951). Each of the above courses of action has an associated payoff or risk which depends on the situation in which the discrimination is to be made. For example, the risk associated with a decision to stop a machine which is involved in an assembly line is greater than is that associated

with the decision to stop a machine tool not integral to a process involving many processing steps. The decision to stop the machine is here assumed to be the result of an aural indication of abnormal performance. The listener's knowledge of the probability of occurrence of signal patterns also is significant.

Formally, whenever the a priori probabilities of the signal patterns are known, and a cost can be associated with the available courses of action, the Bayes criterion can be used to minimize the total risk (Van Trees, 1968). For the sequential hypothesis testing problem, the significant parameters are:

a. P_1, P_2, the a priori probabilities of ω_1, ω_2, respectively.

b. C_{ij}, $i, j = 1, 2$ where C_{ij} is the cost of deciding in favor of signal pattern ω_i when signal pattern ω_j is correct.

c. C_{def}, the cost of deferring the decision.

Alternately, if the a priori probabilities are not known, or if it is not reasonable to assign costs to each of the courses of action, it may be more appropriate to set a criterion which reduces the probability of a particular error below some limit. The threshold adopted in this case will be chosen using the Neyman-Pearson criterion.

Whatever the criterion by which the thresholds for the available courses of action are determined, the observed performance will be strongly influenced by this threshold. A way to characterize this performance under various criteria is by means of a receiver

operating characteristic (ROC) curve (Egan, 1975). This curve presents the probability of a detection P(D) vs the probability of a false alarm P(FA). As the listener adjusts his criteria, both of these probabilities will change.

Under a lax criterion, i.e., say ω_1 whenever there is any evidence of the presence of feature Ω_1, the number of correct discriminations will be high as will the probability of a false classification. A more strict criterion will decrease both of these probabilities.

3.2 Factors Influencing the Terminal Decision

Of the possible decisions in the sequential classification task, a decision to accept hypothesis H_0 and H_1 is referred to as a terminal decision. The way such a terminal decision is made in a detection problem has been treated elsewhere and will not be reiterated (Loeb and Binford, 1970). The concept is extended to a classification task below.

Consider the sequential classification as consisting of a series of hypothesis tests corresponding to dichotomous features. Each such test constitutes a detection opportunity for feature Ω_1. If the hypothesis H_0 and H_1 both have finite probability of occurrence, it is only a question of time until a terminal decision will occur by chance alone.

Wald (1947) extended the concept of the likelihood ratio test to the sequential problem. The sequential probability ratio test is, at the i'th stage,

$$B \overset{H_1}{<} \lambda_i(\bar{R}) = \prod_{j=1}^{i} \frac{P_{r|H_1}(R_j|H_1)}{P_{r|H_0}(R_j|H_0)} \overset{H_0}{<} A \quad , A < B \quad , \tag{3.1}$$

else defer the decision.

Here A and B are the thresholds for deciding H_0 and H_1, respectively. He goes on to prove that such a test will always terminate, and he derives the expected number of observations for a terminal decision.

The definition of a detection opportunity must be made very precisely if a relationship between signal detection theory and operator performance is to be expected. Furthermore, the sequential probability ratio test assumes that the probability densities under H_0 and H_1 remain unchanged from observation to observation. The actual detection occurs in a situation where the relationship of extraneous noise to sound of interest will normally change as a function of time. The classification rules are also likely to change in a real-world situation. If, for example, the time to come to a decision is very long, the listener may relax the criterion for the sake of ending the uncertainty.

3.3 Forced Response vs Sequential Classification

Classifications which imply only the detection of one dichotomous feature may be conveniently visualized by the probability densities under H_0 and H_1. Whenever the listener must either make a yes-no (YN) or a forced choice, two alternative (2AFC) response, a single threshold divides the decision space into two distinct regions. Figure 2a shows the probability densities and error regions for this

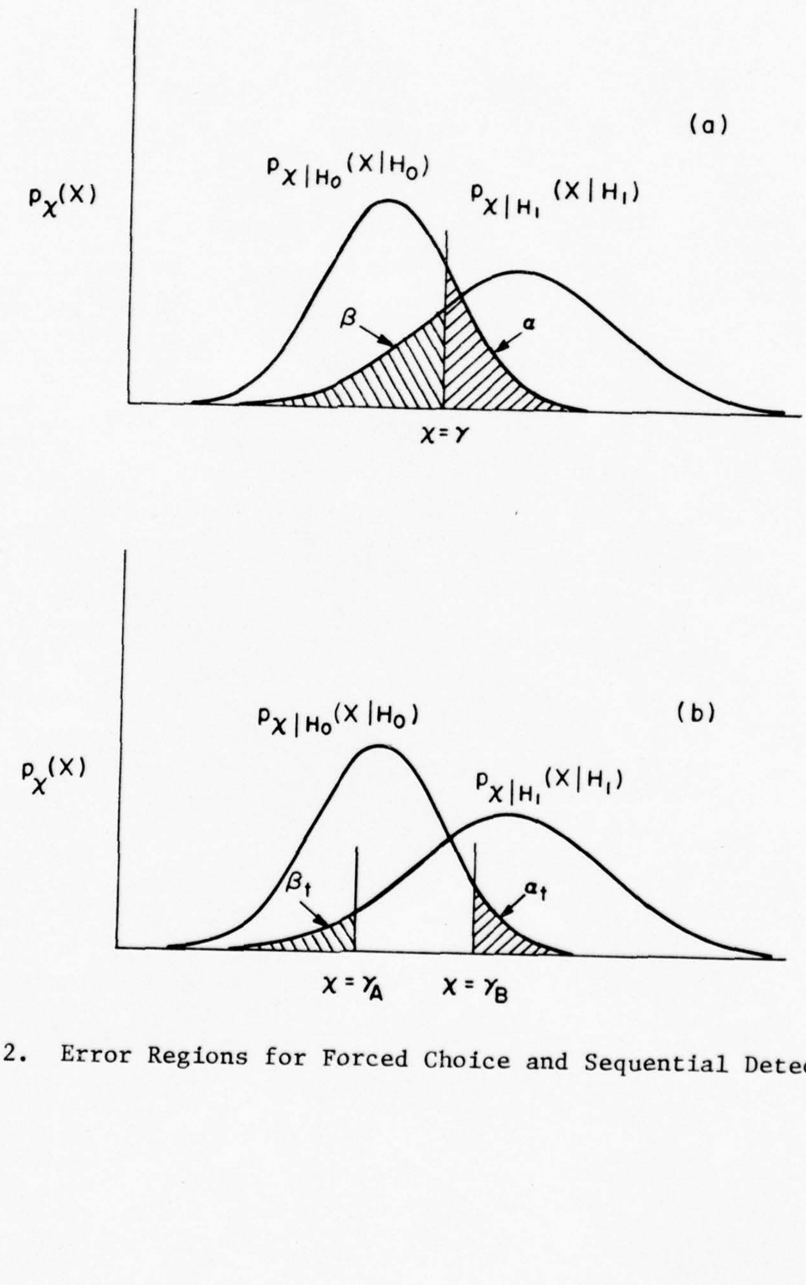


Figure 2. Error Regions for Forced Choice and Sequential Detection.

situation for the detection of a band of noise in noise. The error regions are:

a. $\alpha = P(\text{deciding } H_1|H_0) = P(FA) = \int_{\gamma}^{\infty} p_{X|H_0}(X|H_0)\,dx$ (3.2)

b. $\beta = P(\text{deciding } H_0|H_1) = 1 - P(D) = \int_{0}^{\gamma} p_{X|H_1}(X|H_1)\,dx$. (3.3)

The error regions at time t_i for sequential detection are shown in Figure 2b. The probability of an error in classification will be given by:

$$P(E)_t = P(H_0)\,P(X|H_0)_t + P(H_1)\,P(X|H_1)_t \tag{3.4}$$
$$= \alpha_t\,P(H_0) + \beta_t\,P(H_1) \quad ,$$

where

$$\alpha_t = \int_{\gamma_\beta}^{\infty} p_{X|H_0}(X|H_0)\,dx, \quad \beta_t = \int_{0}^{\gamma_A} p_{X|H_1}(X|H_1)\,dx \quad . \tag{3.5}$$

In practice, the criterion, hence threshold, will be set by the listener in response to the situation. The numeric value of the threshold γ will not be known but can in the forced response case be inferred from the observed values of $P(FA)$ and $P(D)$. The sequential detection and classification case presents some difficulties, however. Especially when the distributions under H_0 and H_1 are changing, knowledge of α_t and β_t is itself not adequate to define γ_A and γ_B without assumptions of the terminal decision process.

The experimental approach described in Chapter IV is designed to elicit information about classification performance when the criterion is assumed constant. Some change in criterion may occur, however, and this aspect is also treated.

CHAPTER IV

DESIGN OF AN EXPERIMENT IN AURAL RECOGNITION

4.1 Introduction

This chapter presents the rationale for the design of an experiment in aural recognition. The experiment incorporates some of the methods used in classical threshold techniques. Also, the data are to be analyzed from the point of view of signal detection theory. The differences between the threshold techniques and experimental methods used by researchers in signal detectability are discussed. The methods are compared and an experimental technique, here called the "modified threshold procedure," is described.

In later sections of the chapter, an extension of the technique to a confidence rating scheme is described. This latter approach is designed to map out the functional relationship between confidence level and operator performance.

The safeguards exercised in experiments using subjects are detailed in Section 4.6. The calibrations and loudness controls are also described in that section.

4.2 Threshold vs Forced Response Tests

Classically, the basic correlates of the auditory stimulus have been investigated by threshold methods (Boring, 1950). In this approach, the stimulus parameter to be measured is changed slowly until a just perceptible difference in the physiological response occurs (Licklider, 1951). One related area where this technique has been used extensively is in the investigation of the loudness of various spectrally shaped noises (Cremer, Plenge, and Schwarzl, 1959; Zwicker, 1958; Zwislocki, 1969).

The results of these threshold experiments can be explained using the theory of signal detectability (Treisman and Watts, 1966). It can be inferred that the listener in such experiments operates in a sequential detection or recognition environment. Whenever the established criterion is exceeded during a period of increasing stimulus intensity, the threshold can be thought to have been exceeded.

It should be pointed out, however, that single or multiple thresholds for sensual inputs have been postulated as explaining the observed results. Also, quantization of the auditory process has been suggested (Licklider, 1951). While this thesis is not concerned with the analysis of the detailed mechanisms which are active in the human, there are fundamental differences in approach between experiments using threshold techniques and those designed to test various aspects of the signal detection theory.

Detection theory related experiments are designed to elicit either data about the ROC curve or the psychometric function of the listener. The psychometric function is the variation of percent

correct detection or classification as a function of stimulus level. The key parameters are usually the various probabilities of the available courses of action and the detectability d'_{opt} expected on the basis of some signal detection model. The detectability is usually defined as the normalized difference of means of a statistic under two alternate hypotheses.

The signal detection theory types of experiments present the stimuli to a listener at some signal-to-noise ratio and observe the performance. The criterion is usually specifically included by way of instructions prior to the test. Data are taken at a number of signal-to-noise ratios or other fixed values of a parameter (Swets, 1964). For YN type experiments, the responses are either "yes" the signal was present in a trial period, or "no" it was not. In forced choice, N alternative (NAFC) experiments, the listener indicates in which one of N trial periods the signal occurred. The signals are usually presented in the presence of an interferring noise at some preselected level.

The YN or NAFC experiments have the advantage that the results are subject to well known and easily applied statistical hypothesis testing techniques as are described in Chapter II (Egan, 1975).

4.3 A Pilot Study and Its Impact

Initially, a pilot study was conducted using University students as subjects to identify weaknesses in experimental techniques and to obtain bounds on listener performance. This was an identification task in which the listener was asked to decide which of two marine

sounds (the exposure set) was subsequently presented in a broadband noise background. The signals were presented binaurally from tape recordings in an experimental arrangement very similar to that which is described in later sections. However, for these initial tests, the unknown or probe stimulus was presented for a 20-second period of time at a fixed signal-to-noise ratio. The listener was asked to record on an answer sheet his or her assessment of which member of the exposure set the probe corresponded to. In addition, the third alternative "don't know" was to be indicated whenever "reasonable" doubt existed.

Because the sounds presented do not fall within the range of normal auditory experience for these listeners, an immediate area of difficulty arises. One is precluded from using techniques applicable to, for example, speech intelligibility studies or other recognition tasks for meaningful material where the subjects can be assumed to have a common learned ability to distinguish between stimuli (Pollack, 1948; Olach, 1964). A time consuming stimulus training period is required prior to any probe measurement. Since the stimuli characteristics cannot be reliably verbalized, an arbitrary designation of A and B were associated with the trained stimuli. Such arbitrary association of letters with the members of the exposure set is unfortunate. However, the method used by Pollack (1959) to circumvent this shortcoming, the method of recognition memory, cannot be used in the present case. In the method of recognition memory, the probe is chosen from an augmented set which includes the members

of the exposure set and members of another set, the confusion set. The listener response then consists of an assessment of whether the probe in fact corresponded to a member of the exposure set. The marine sounds are such that the experimenter cannot know if the confusion signals will, in fact, sound "like" one of the exposure signals.

A test event consisted of an exposure (learning) period for the two signals without interfering noise. After a short pause, the probe was presented in the interfering noise. Subsequently, the exposure set was repeated in a refresh period, or a new set of signals was exposed for learning. No feedback was used except that in some cases the subjects were appraised of their performance at the end of a test session. Each test event requires about two minutes to complete for a total of 15 to 18 events per session.

Little was known at the time of this pilot study about the range of signal-to-noise ratios at which the listener performance would attain some pre-determined probability of correct response P(C). It was necessary, therefore, to present the probe at a number of values of signal-to-noise ratios. Using four signal pairs of interest and five values of signal-to-noise ratio, there were 20 performance indices to be estimated. As will be shown later, at least 50 data points are desired for each performance index for a total of 1000 events or about 60 sessions.

The availability of trained listeners was limited, and not all performance indices were adequately studied. This method was abandoned in favor of the ramped signal-to-noise ratio technique

discussed in subsequent sections of this report primarily because a faster method was required. This early study was very useful in identifying several weaknesses in procedures and equipment design.

4.4 A Modified Threshold Procedure

Most of the results presented in this report were collected using a modification of threshold procedures. The scheme used is modeled after the techniques used by Cremer et al. (1959) and Zwicker (1958) to determine the loudness of various bands of noise. Dubrovskii and Tumarkina (1967) also used an experimental technique which is very similar to the one described below. Some considerations which strongly influenced the experimental technique used in this thesis were:

a. Because it was desired at the outset to make the listener related portion of the system portable, the signals are pre-recorded on magnetic tape. Also, the response recorder is of necessity simple.

b. It was desired to automatically record responses without the test conductor needing to be present. Subjects were thus able to conduct tests at their own discretion with a minimum of scheduling difficulties.

c. The process of data reduction was to be a simple one with a minimum of manual data handling.

In the modified threshold procedure, the probe is initially presented in a broadband noise background at a low signal-to-noise ratio. This signal-to-noise ratio (SNR) is then increased step-wise with time until such time as the listener can make a determination of which exposure signal the probe corresponds to. This method has an advantage in its similarity to the sequential classification task encountered in many real situations. However, the results only provide data about the probability of correctly classifying at the SNR needed for a high confidence terminal decision.

The sequence of occurrences during each test event is shown in Figure 3. The exposure signals A and B are chosen with randomization from the set of auditory patterns of interest. Generally, these patterns differ by one or more features. The initial and refresh exposures are of fixed duration which the listeners generally agreed on as sufficient for later recognition. The probe SNR is always low initially, but this value is randomized to avoid listeners responding on the basis of time instead of the perceived aural response.

Responses were indicated by listeners by pressing either a switch marked A or one marked B. Once the listener indicated a decision, the auditory signal was blanked for the remainder of the response period. The duration of the response period was also randomized. By not knowing when the response period would end, the listeners were in effect assigned a cost for continuing in the listening state. It was found that subject performance could be established by means of verbal instructions prior to each test session. Once trained, the subjects maintained a consistent criterion

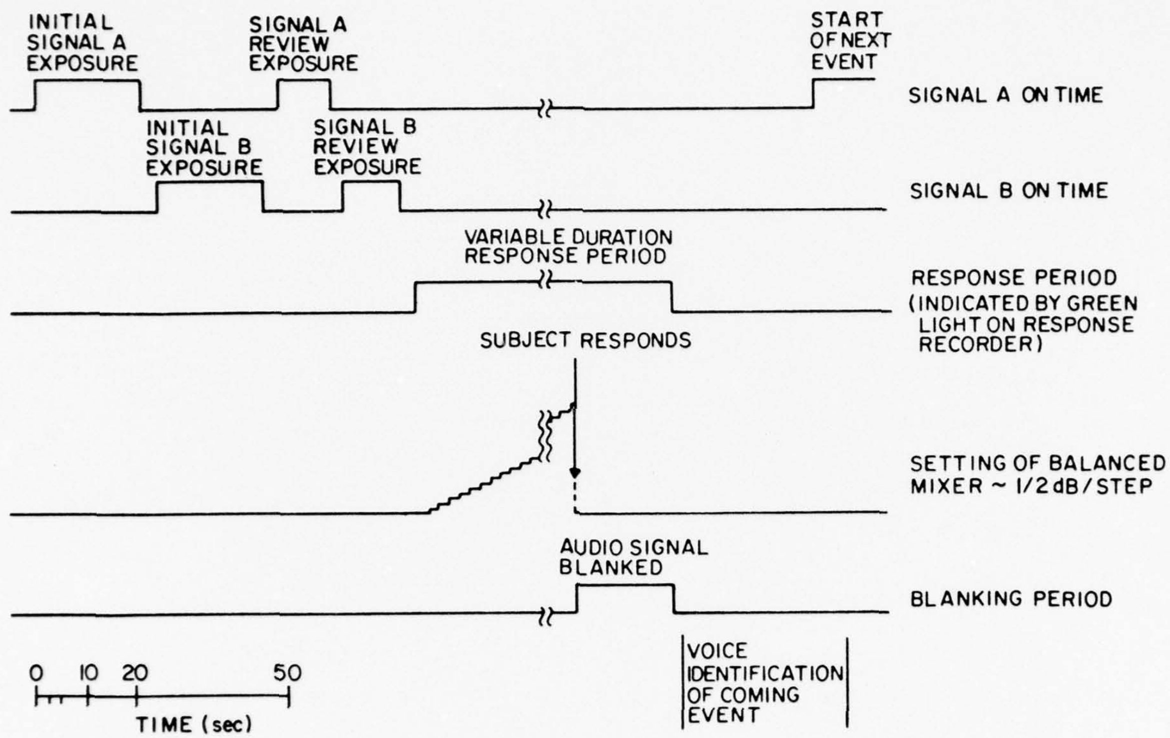


Figure 3. Timing Diagram for an Event in the Modified Threshold Procedure.

which was described in the instructions as "reasonably certain."

The pre-recorded signals were played back using a Crown 700 tape player and a pair of Telephonics Model TDH-39 dynamic earphones. The signal was fed in phase to a matched set of phones installed in an audiometric booth where practical. Tests done on-site using Navy personnel were conducted in the quietest classroom available. Figure 4 diagrams the equipment used at the listener location. The dual channel audio source tapes contain a pilot channel for response recorder control and a channel with the stimulus signal. This latter tape channel is recorded with the exposure, probe, and interfering noise signals. Control signals recorded on the pilot channel consist of a 12.5 kHz reference tone during the exposure period and a tone with frequency proportional to SNR during the response period. These control signals are decoded by means of a phase locked loop set to 12.5 kHz and a level detector. Digital logic is then used to control the state of the response recorder and of a cassette recorder.

A Sony cassette recorder is used to log responses which are coded as frequencies. The tone proportional to SNR is also output to the cassette during the response period prior to the subject's responding. During the exposure period, the cassette recorder is stopped so that a session which lasts nearly an hour requires considerably less cassette time. The equipment at the listening site is pictured in Figure 5.

Data were analyzed in two ways. Initially, the cassette tapes were played back on another cassette tape unit and output on a cash register tape using a printing frequency counter. The SNR at the

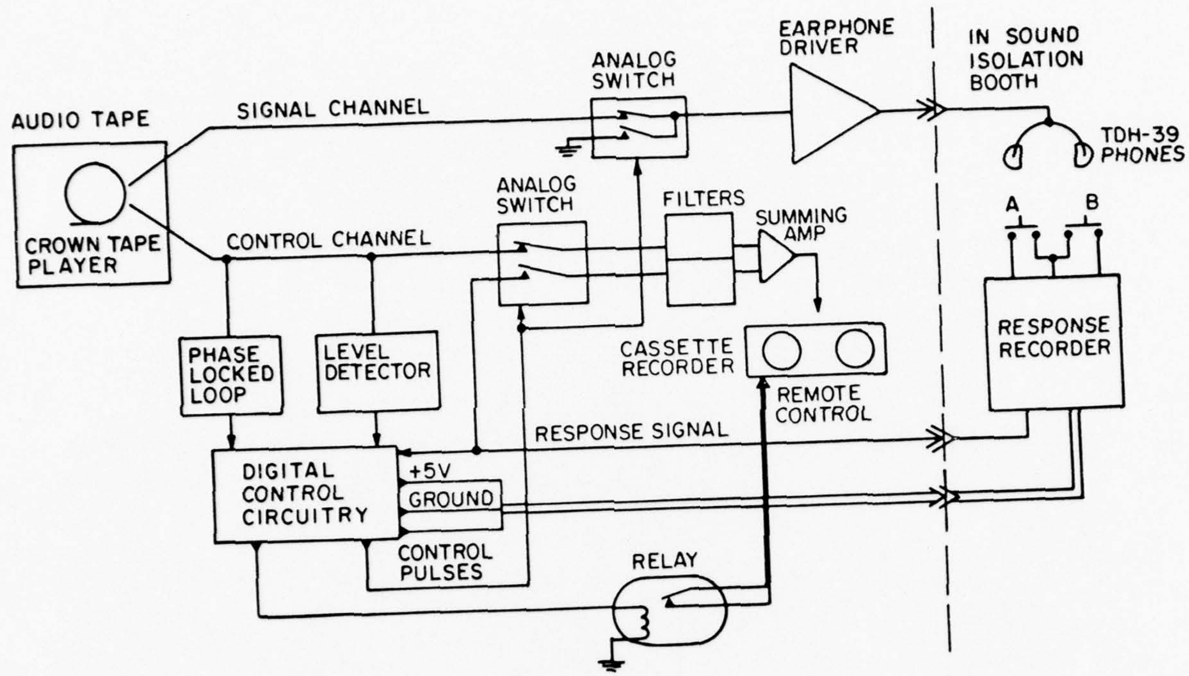


Figure 4. Summary of Equipment Used at the Test Site for Modified Threshold Experiments.

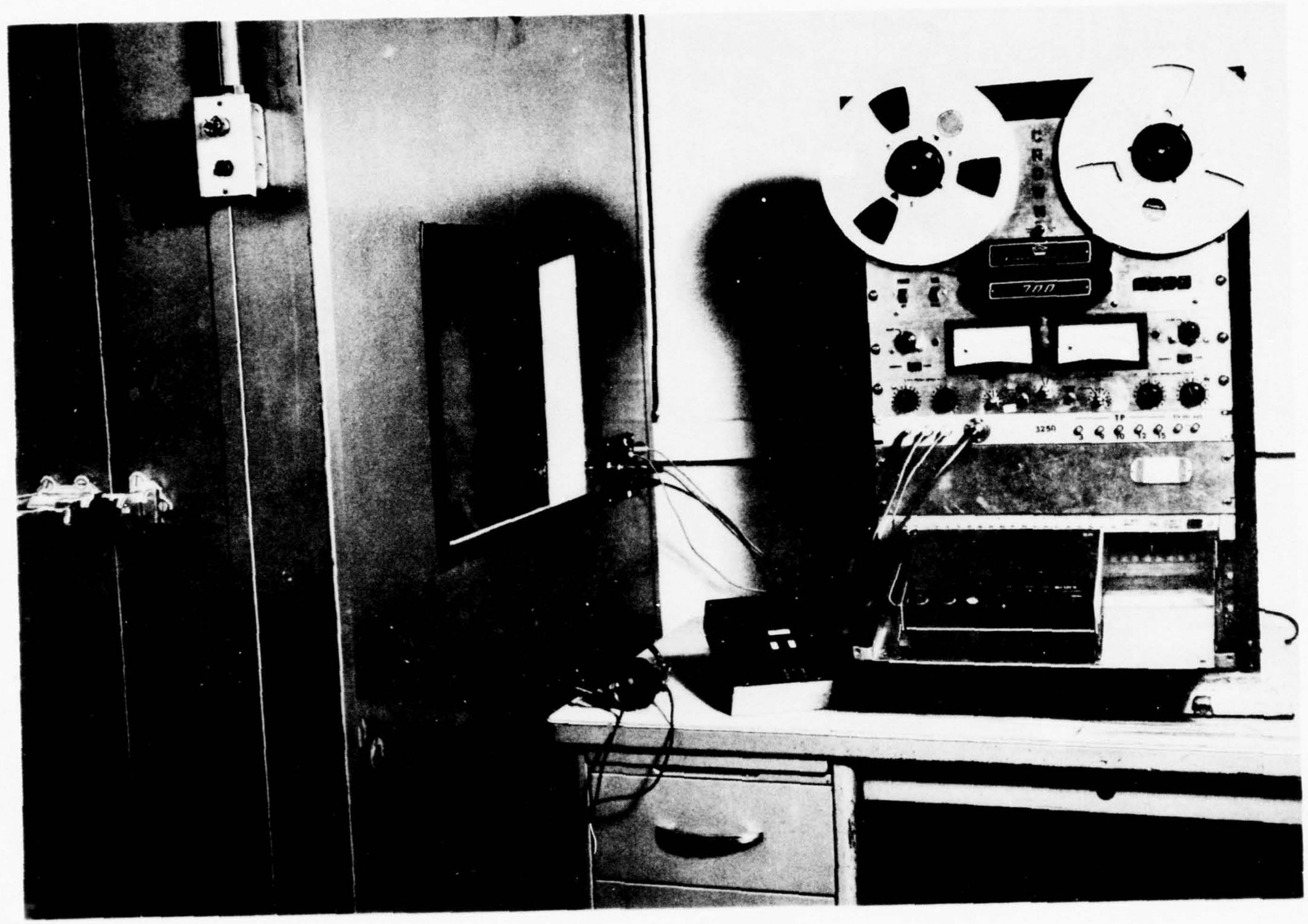


Figure 5. Photograph of Audio Tape Playback System with A/B Response Recorder.

moment of subject response and the classification (A or B) were determined manually. As the number of data points increased, the results were key punched and copied onto a computer magnetic tape for retention and subsequent data reduction. Later, the process was automated by outputting the counter frequency directly to phase encoded digital tapes. A buffered Pertec digital tape unit was used for this purpose. These IBM compatible digital tapes are amenable to direct processing by computer.

4.5 A Continuous Rating Experiment

The terminal decision logged in the above method does not give enough information to allow separation of criterion effects from those due to the detectability of the stimulus. In order to obtain information of how performance shifts with criterion, a continuous rating experiment was also undertaken. In this approach, the subject continuously indicates his tentative classification (A or B) as well as his degree of certainty. The responses are indicated by moving a linear potentiometer control from side to side. The potentiometer positions are annotated with a numeric scale ranging from 0 to 10 as is shown below.

A	B
10 8 6 4 2 0	2 4 6 8 10

The listeners were instructed to attempt a preliminary, albeit low confidence classification as soon as possible and to modify this decision as more data became available. As the SNR increased, they would tend to move up the scale to higher confidence levels. These experiments are hereafter referred to as "slide wire" tests.

For the slide wire tests, a number of subjects were tested at once. Up to five response recorders were set up and each subject location was equipped with its own response recorder and a pair of Pioneer SE205 circumaural earphones. Responses were output to a six-channel strip-chart recorder along with a voltage proportional to SNR. This latter voltage was derived from the pilot channel of the dual channel audio tape. The SNR proportional frequency was converted to a voltage by means of a frequency to voltage converter. The subject related portions of the equipment used in slide wire tests is pictured in Figure 6.

The data on the strip-chart recorder output were manually reduced to obtain the confidence and classification decision at various signal-to-noise ratios. These data were also ultimately reduced to computer compatible form.

4.6 Listener Safeguards and Calibration Techniques

Prior to beginning experiments using human subjects, approval was obtained in compliance with the Institutional Assurance provisions of The Pennsylvania State University. For approval, a prospectus of the study was submitted. This proposed approach was reviewed by a University Committee to ascertain that no danger to the listeners would result. Specifically, the levels and durations of signals are such that no damage risk is incurred (American Academy of Opthalmology, 1957). Each subject was instructed as to his role as participant in the study, and it was pointed out that under no conditions would the sounds be so loud as to cause discomfort.

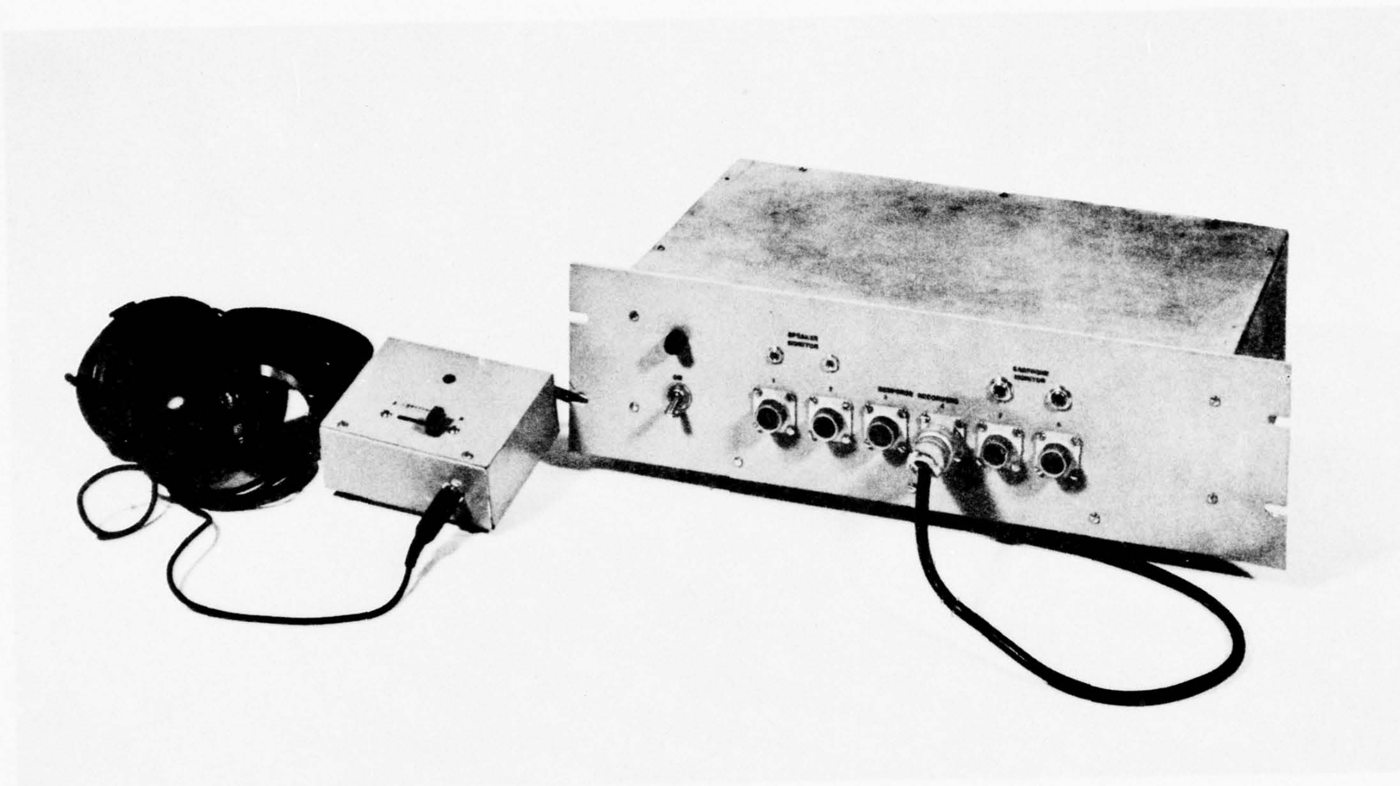


Figure 6. Photograph of Slide-Wire Response Recorder and Associated Equipment.

See also Appendix A.

In order to eliminate the effect of shift in subjective loudness as the probe signal to interfering noise ratio changed, a balanced mixer followed by an automatic loudness control was used. The loudness as presented to the subject was carefully maintained at a loudness level of 65 phons (GD) (ISO R532). This loudness level was verified from 1/3-octave band measurements of the voltage function to the headphones and taking into account the factory provided earphone calibration with the MX41/AR ear cushions. This loudness level applied to the exposure signals as well as the probe signal in noise. The automatic loudness control used was a relatively simple one consisting of an automatic level control with B weighting in the feedback (Peterson and Gross, 1963). The marine sounds used were simple enough so that there was no need to use complex circuits for control of loudness as are needed in general sound broadcast work (Bauer and Torick, 1966).

Since it was not possible to directly calibrate the circumaural earphones used in the slide wire experiments, the loudness was adjusted using a loudness balance technique. For this adjustment, the same noise signal was fed to one of the TDH-39 calibrated phones and to one of the Pioneer phones and the level to the latter phone was adjusted until the reference signal on the left ear sounded equally as loud as that in the uncalibrated phone on the right ear. The positions were then interchanged. This calibration is not especially accurate because it is difficult to maintain proper

pressure on the TDH-39, and also the circumaural phones are sensitive to the shape of the pinna and the subject's hairline. However, the absolute calibration need not be very precise for the following reasons:

a. The loudness is quite constant and sufficiently high so that all signals are well above the auditory threshold and where the ear is reasonably linear.

b. Since there is no discomfort, and the calibration is probably no more than 10 phons in error, the damage risk to the listeners is never significant.

c. Differences between units of the TDH-39 and Pioneer earphones in frequency response will tend to be at very low and high frequencies. These are frequencies which do not affect the dichotomous feature perception for the sounds of interest.

The SNR was controlled by means of a precision balanced mixer with step increments of about 1/2 dB. The shaft of this mixer is connected through a reduction gear to a potentiometer which varies the voltage to a voltage controlled oscillator. This frequency was measured at each step of the mixer and used as input to subsequent data reduction. The combined frequency uncertainty due to flutter and wow in the Crown tape player, the Sony cassette recorder, and the final cassette player was less than $\pm$ 1/4 of the frequency shift resulting from a change of SNR of 1/2 dB.

Each of the two exposure signals was analyzed and plotted against the interfering noise background using 1/3 octave analysis accurate to 1/2 dB. This calibration was performed with the balanced mixer set to zero. The linearity of the balanced mixer was checked with two widely spaced tones and the relationship of mixer setting to resulting SNR was determined and served as a correction in subsequent processing.

Using these calibrations and correction terms, the SNR output to the tape recorder used for recording the primary tapes is therefore known for all combinations of background noise and balanced mixer settings. Subsequent recording steps were also carefully monitored to insure that the dynamic range of the tape medium was never exceeded. These steps then insure that the effective SNR presented to subjects could be accurately determined. The calibration and correction curve accuracies are such that SNR's can be determined to within 1/2 dB.

CHAPTER V

EXPERIMENTS IN AUDITORY PATTERN DISCRIMINATION

5.1 Summary of Experiments

The two techniques described in the previous chapter, modified threshold and slide wire, were used to measure the auditory pattern discrimination for noise-like sounds. Experiments were conducted over a period of about a year using trained University student listeners as well as submarine sonar operators with various degrees of experience.

This chapter describes the experiments in detail as to signals used, statistical methods applied to data reduction, and subject training. The rationale for choosing the signals used in the discrimination tasks is presented in Section 5.2 along with characterizations of the marine sounds. Section 5.3 presents the statistical testing techniques used and then, in Section 5.4, the steps taken to train the listeners in this paradigm are described. It is argued that consistent performance was attained with the two populations of listeners used.

The student subjects only participated in the modified threshold experiments. For student tests, the signals were presented under calibrated earphones in an audiometric booth located in the Applied Science Building of The Pennsylvania State University.

Seven students were involved in those tests and they were concurrently, or had been, participants in other psychoacoustic studies at the University. There were three advanced standing female students and four graduate school males with normal hearing as verified by current audiograms. These listeners were free to conduct tests at their own convenience with the following restrictions:

a. No two test sessions could run sequentially.

b. No more than two test sessions could be held on one day.

The students averaged one test session per week.

Submarine sonar operators participated in both slide wire and modified threshold experiments. These tests were conducted in the quietest classroom available at the US Naval Submarine Training Center, Pacific, San Diego detachment. These operators took part in one session per day for five days. Two such week-long tests were conducted using a total of 12 sonar operators with between 6 months and 20 years experience. During these tests, the author or an assistant were present at all times when tests were in session. All of these operators had been trained for passive submarine sonar tasks and all had some experience in that setting. Those operators with extensive ship-board experience were at the time serving as instructors at the training center.

The calibrated TDH-39 phones were used with sonar operators for modified threshold tests. Although attempts were made to find a quiet location for the tests, some interfering sounds in the room resulted from vehicular traffic outside. Extensive notes were taken

throughout the tests of such interferences and suspect events were eliminated from the data later. Several operators participated in the slide wire experiments at the same time. The Pioneer phones discussed in Chapter IV were used for these tests.

5.2 Choice of Noise-Like Signals

The objective of this work was to quantify subjective responses to sounds, especially in the discrimination between similar noise-like stimuli. To achieve this objective and to allow the use of sounds which play a significant role in an everyday situation of interest, marine sounds with very well defined characteristics were chosen. Since submarine sonar operators are routinely required to distinguish between marine sounds, this is one area where discrimination ability is known to be important. Furthermore, the trained sonar operators constitute a population of listeners with hopefully comparable training and similar work-a-day environments. The marine sounds are also ones for which discrimination can be easily trained with experienced listeners as are the University students used.

Many industrial noises have characteristics which are similar to the marine sounds. In fact, the marine sources exhibit broadband noise spectra which are caused by the same mechanisms leading to many sounds in industry, principally hydrodynamic sources and the composite of many independent vibrating components. Kinsler and Frey (1962) describe the radiated sound of ships as resulting from "cavitation," a hydrodynamic mechanism, and from "machinery," a source common to most industrial situations.

The stimuli used in the present tests were of two types. The first type of sounds to discriminate were actual marine sounds which were contrasted with the same sound modified by removal of selected features. In addition, there were sounds generated from stationary noise sources to have controlled characteristics resembling those of the marine sources. Two such customized sounds were then contrasted in the discrimination tests. The two laboratory-generated sounds used in such tests differed by one or more dichotomous features.

The sound patterns ω_1, ω_2, and ω_3 associated with the unmodified marine recordings are characterized in Table 1. These sounds were from magnetic tape recordings and were processed by a whitening filter to reduce the overall dynamic range. The whitening filter removes the approximately 6 dB per octave decrease in sound pressure level with increasing frequency as is exhibited by many marine sounds (Kinsler and Frey, 1962). See Appendix B for a detailed description of equipment used to process the raw data and for generation of the final audio tapes.

The dichotomous features tested were the result of modifications of the original marine recordings. These modifications consisted of filtering or removing of amplitude modulation by using a high speed automatic gain control. The resulting signal patterns are listed below:

a. High pass filtering with a sharp filter with 3 dB down point at 707 Hz to conform to the edge of octave 30 results in signal patterns identified as ω_{1H30}, ω_{2H30}, and ω_{3H30}.

TABLE 1

SUMMARY OF SOUNDS ASSOCIATED WITH SOUND PATTERNS ω_1, ω_2, AND ω_3, MARINE SOUNDS

Sound Pattern	Description of Its Sound
ω_1	High speed rhythmic sound with pronounced bursts of high frequencies but of short duration.
ω_2	Very regular rhythm like that of a steam engine or slow train. Some irregular high frequency popping.
ω_3	Noise-like without any noticeable rhythm. Sounds like that of high speed air but at the same time having a low rumbling component.

b. High pass filtering with a sharp filter with cutoff frequency of 354 Hz to conform to the edge of octave band 27 results in signal patterns ω_{2H27} and ω_{3H27}.

c. Elimination of temporal amplitude fluctuations slower than about 70 Hz results in signal patterns ω_{1SN} and ω_{2SN}. Here the suffix SN is used to indicate that the resulting pattern is approximately a stationary noise.

Table 2 summarizes the signal pattern pairs tested with either student or sonar operator subjects. The number of data points with each group shown in Table 2 is the number of useful events after training events and those for which results are suspect due to extraneous noise interference or response recorder malfunctions have been eliminated. This total does, however, include data collected using either modified threshold or slide wire techniques. The tabular headings $\omega{:}H_0$ and $\omega{:}H_1$ are used to indicate the signal pattern with the dichotomous features absent and present, respectively.

The reasons for choosing these modified characteristics of noise-like sounds are treated in greater detail in Chapters VI and VII. The choice of filter cutoff frequencies, however, was made so as to conform to recommended octave bands as detailed in USA Standard S1.6-1967 to insure repeatability.

Customized signals were made by filtering stationary broadband noise and by modulating the noise in some cases. These signals were designed to exhibit many of the characteristics of the marine sounds but with precisely defined properties. Table 3 describes signal

TABLE 2

SUMMARY OF EXPERIMENTS USING
TAPE RECORDED MARINE SOUNDS

Experiment	Purpose	Signal Patterns		No. of Data Points	
		$\omega:H_0$	$\omega:H_1$	Student	Sonar Operator
I	To test the discrimination ability for high pass filtered noise-like sounds with cutoff frequency 707 Hz.	ω_{1H30}	ω_1	70	17
II	To test the discrimination ability for high pass filtered noise-like sounds with cutoff frequency 707 Hz.	ω_{2H30}	ω_2	82	82
III	To test the discrimination ability for high pass filtered noise-like sounds with cutoff frequency 707 Hz.	ω_{3H30}	ω_3	124	140
IV	To test the discrimination ability for high pass filtered noise-like sounds with cutoff frequency 354 Hz.	ω_{2H27}	ω_2	-	18
V	To test the discrimination ability for marine sounds which have been treated to remove amplitude modulation.	ω_{1SN}	ω_1	37	47
VI	To test the discrimination ability for marine sounds which have been treated to remove amplitude modulation.	ω_{2SN}	ω_2	-	53

TABLE 3

SUMMARY OF LABORATORY GENERATED
SOUND PATTERNS $\omega_4 - \omega_9$

Sound Pattern	Description of The Sound
ω_4	A sound consisting of two octave bands of noise, one centered at 500 Hz (Band 27) and one centered at 4000 Hz (Band 36) where the low frequency band is amplitude modulated with a 10 Hz square wave.
ω_5	Same as ω_4 but the low frequency octave band is a stationary noise without temporal changes in amplitude.
ω_6	Same as ω_5 but with the low frequency octave band having reduced amplitude and hence detectability.
ω_7	A sound consisting of a single octave band of noise centered at 4000 Hz (Band 36).
ω_8	Same as ω_5 but with a 10 Hz square wave modulation of the high frequency band (Band 36).
ω_9	A single octave band of noise centered at 4000 Hz but with 10 Hz square wave modulation.

patterns ω_4 to ω_9. The experiments performed using the customized signals are summarized in Table 4. None of these experiments was done using student subjects.

The background or interfering noise against which the unknown or probe signal was presented was an ocean ambient for signal patterns ω_1, ω_2, and ω_3 (Perrone, 1970). This ocean ambient was the result of 1/3-octave filters summed with appropriate weighting and exhibited a nearly flat spectrum with somewhat accentuated low frequency components. The background noise used with signal patterns ω_4 to ω_9 was a white noise with the very low frequencies filtered out to avoid audio tape saturation. Figure 7 shows the 1/3-octave spectra of the background noises used. Note that a spectrum which is white, i.e., constant energy per Hz across frequency, will have a 1/3-octave spectrum which increases at the rate of 3 dB per octave because of the bandwidth proportionality exhibited by such spectrum analyzers (Peterson and Gross, 1963).

5.3 Analysis of the Data

The measured quantities resulting from the modified threshold experiments were, after applying corrections and calibration factors:

a. The signal-to-noise ratio (SNR) needed for the subjects to commit to a terminal decision of "reasonable certainty."

b. The classification $\hat{\omega}(R)$ corresponding to the exposure signals A or B.

c. Time from onset of the response period to the time when the terminal decision was made.

TABLE 4

SUMMARY OF EXPERIMENTS USING LABORATORY
GENERATED SIGNAL PATTERNS ω_4, ω_5, ω_6, or ω_7.

Experiment	Purpose	Signal Patterns		No. of Data Points
		$\omega:H_0$	$\omega:H_1$	Sonar Operators
VII	To evaluate the discrimination performance for amplitude modulation.	ω_5	ω_4	53
VIII	To evaluate discrimination for a stationary band of noise.	ω_7	ω_5	55
IX	To evaluate discrimination for a stationary band of noise, reduced detectability.	ω_7	ω_6	46
X	To evaluate discrimination in the presence of amplitude modulation.	ω_9	ω_8	23

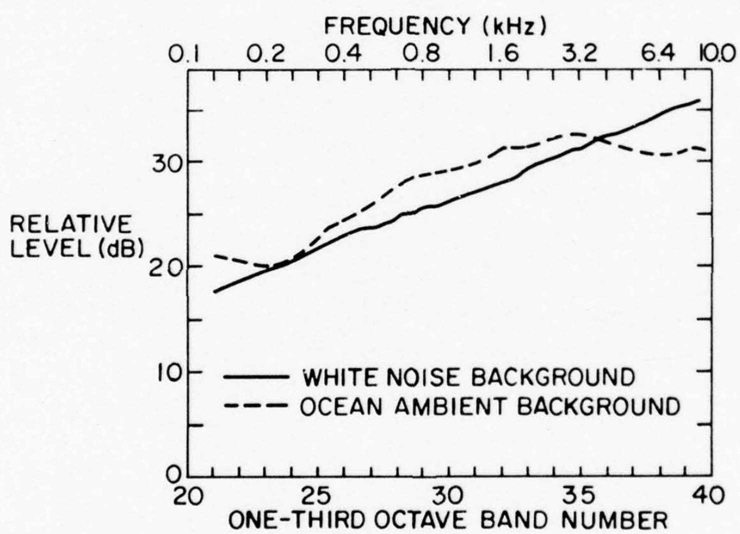


Figure 7. One-Third Octave Spectra for Background Noise Against Which Probe Stimuli Were Presented.

Of these experimental measurables, the SNR requires additional explanation. The SNR has been variously described as the relative energy over a band of frequencies, as the ratio of background to signal spectral level at one frequency, or as the ratio of integrated signal and background noise powers (Gundy, 1961; Stallard and Leslie, 1974). In this thesis, the SNR serves as a reference only and is defined as the difference in dB of the 1/3-octave spectrum level of the signal (desirable noise-like sound) and the background noise against which the probe is presented. Furthermore, this difference in dB is measured in the 1/3-octave where the signal has the highest spectral level relative the background. Without exception, this reference 1/3-octave band falls at higher frequencies, above 2 kHz, for the signals used.

Because of the step-wise increase of SNR with time, the SNR of the probe relative background remains constant for about two seconds between changes. However, the subjects were unable to determine if the SNR was increased linearly or in steps. There were no transients which could have indicated when the steps occurred.

The importance of the time of response is related to possible short term memory effects or due to the effect of shadowing (Klatzky, 1975; Parkinson, Parks and Kroll, 1971). Also, the listener may find it necessary to shift the criterion for a decision as time progresses. In later chapters, the importance of criterion will be addressed in relation to response uncertainty and confidence ratings.

To analyze the experimental data, the effect of applying different treatments to various populations of subjects was investigated. The treatments consist of particular exposure set and probe combinations. Hence, a situation where the signal pattern ω_1 is associated with exposure signal A, signal pattern ω_2 with exposure signal B, and signal pattern ω_1 is used as the probe signal in an ocean ambient noise, is a specific treatment. The results of several treatments may, however, be combined (pooled) in some of the analyses. The populations consist of groups of listeners for whom the discrimination performance is expected to be similar by virtue of training or background. Analysis of the data shows that there are essentially three populations involved in these tests:

a. Student listeners exhibited comparable performance and all seven were found to apply criterion in a similar fashion.

b. Submarine sonar operators with less than two years experience tended to show similar performance.

c. Submarine sonar operators with more than two years experience comprised a distinct population. They seemed to apply criteria differently from members of the other two populations.

Other than partitioning the results according to these populations, no distinction is made between subjects in a group. This method of analysis is tantamount to assuming that there are no individual differences for subjects within a population (Murdock, 1968).

The numerical analysis of the data consisted of computing pertinent first order statistics for treatments applied to populations and of statistical hypothesis tests used to determine differences due to treatments or populations. The principal tests were:

a. Calculation of the mean SNR to respond and the determination of confidence limits on the basis of the sample variance.

b. Determination of the probability of a correct response P(C) and the confidence interval of the estimate assuming that there was an equal probability of correct response for each comparable event.

c. Computation of the T statistic and application of the t test to determine if there is a significant difference between mean SNR ($\overline{SNR}$) as the result of treatments or populations.

To apply the t test for difference of means, fifteen to twenty events are sufficient for most tests (Walpole and Myers, 1972). This is due to the fact that the t distribution quickly stabilizes as the number of degrees of freedom is increased.

However, a large number of events is needed to accurately estimate P(C). Assuming that events comprise Bernoulli trials with equal probability of occurrence, the binomial distribution

$$f(x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0,1,2 \ldots n \qquad (5.1)$$
$$= 0 \text{ elsewhere}$$

applies for x observed correct classifications in n trials (Hogg and Craig, 1970). Murdock (1968) shows that this binomial distribution may be transformed to an approximately Gaussian distribution using the transformation

$$\theta = 2 \text{ arc sin } \sqrt{x/n} \quad . \tag{5.2}$$

The variance of the transformed variable will be:

$$\sigma_\theta^2 = 1/n \quad . \tag{5.3}$$

Hence, θ is $n(\theta_0, 1/n)$ where θ_0 is the transformed sample probability of correct classification.

The 90 percent confidence interval on the probability of correct classification is given approximately by:

$$[\theta_0 - Z(0.95)/\sqrt{n} < \mu_0 < \theta_0 + Z(0.95)/\sqrt{n}\ ] \quad , \tag{5.4}$$

where $Z(0.95)$ is the Z score corresponding to the one tailed normal distribution (Murdock, 1968). Also,

$$Z(0.95) = 1.69 \quad .$$

Figure 8 shows the behavior of this confidence interval for the two cases where the actual probabilities of the occurrence of a correct decision are 0.7 and 0.5, respectively. It is seen that the number of events must exceed about seventy for the 90 percent confidence interval to be within $\pm$ 10 percent of the correct value. The number of events needed to estimate this parameter is therefore

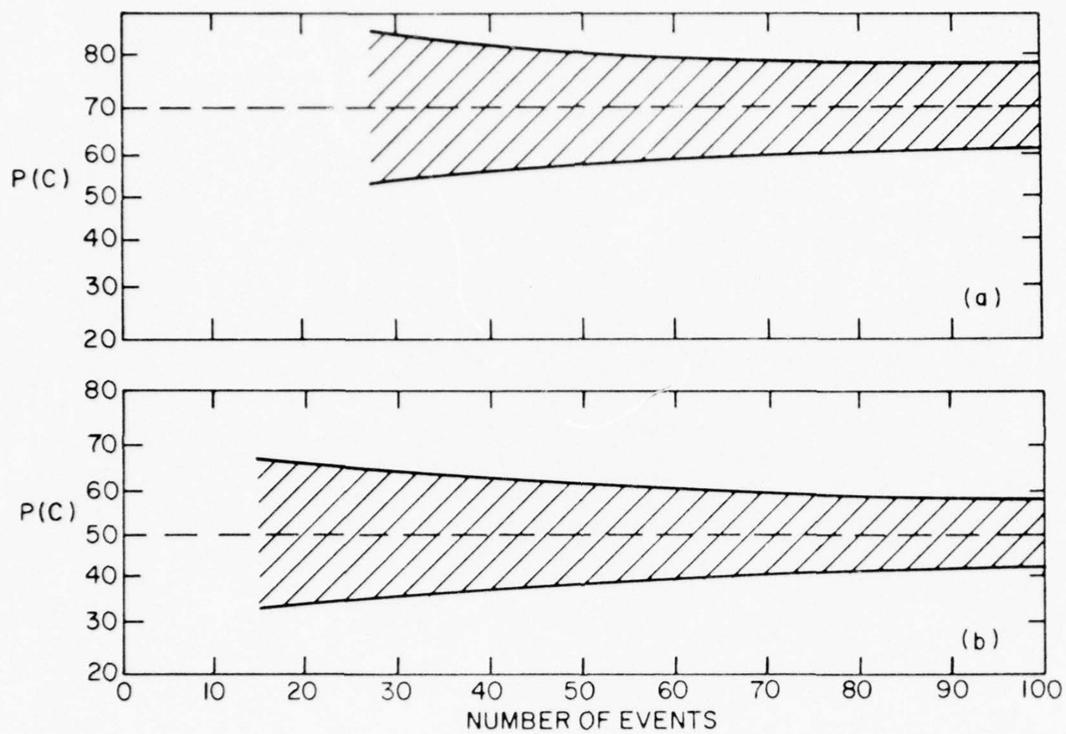


Figure 8. Confidence Intervals on P(C) as a Function of Number of Events.

greater than 50 and preferably 70 if the confidence interval is to be at all useful. Workers using the forced response techniques routinely use 300 or more events to estimate each data point (Swets, 1964; Green and Swets, 1966; Robinson and Watson, 1972).

Hence, whenever the performance index P(C) is to be accurately estimated, many events must be tested. However, tests for difference in SNR due to treatments or populations can be done with a smaller number of events. This fact was taken into account in the design of the individual experiments.

5.4 Subject Training

The training of subjects for participation in the experiments was organized to standardize the subject's criterion. However, some improvement in performance would be anticipated as the signal patterns become more familiar.

The student subjects were payed a fixed amount for each session they participated in. Prior to starting, they were given general instructions about the objective of the experiment and the ground rules. Appendix A documents the basic instructions given as well as the tape recorded instructions. These latter instructions are recorded at the beginning of each audio tape used for these studies and, although they differ somewhat for sonar operators, stress the salient features of experiments for all subjects.

The students participated in pilot studies prior to beginning sessions using the modified threshold procedure and also were exposed to a minimum of five sessions for which the response data was not

processed. Feedback was given only when a subject was obviously applying criteria differently than the norm, or if there was some question about proper understanding of the procedures. When the SNR to respond early in the subject's involvement was compared to the SNR to respond to the same treatments after extensive experience, no significant change was found.

An independent test using a naive subject demonstrated that there was no further improvement in performance after ten sessions, fewer than conducted by the student subjects prior to beginning data collection. In fact, the naive subject showed a dramatic change in performance over only the first five sessions.

Training of the sonar operators consisted of familiarizing these listeners with the experimental procedure. It was assumed that the sonar training would have established their ability to distinguish sounds at a high level. These operators were given an initial interview at which the general objectives of the experiment were outlined. Then, prior to beginning tests with the modified threshold (A/B) response recorder, they participated in a session using the slide wire technique. This exposure gave them an opportunity to listen to the sounds over a large range of SNRs and to ask questions.

The first session with the A/B response recorder was a short one designed to allow for questions by the sonar operators. This data was not used. Application of the t test for change of $\overline{SNR}$ between sessions two and five exhibited no significant change as their participation in the experiment continued.

In order to investigate more treatments in the available time, groups of six events used the same exposure signals but did include randomization of probe signals. This seemingly minor revision of experimental method between students and sonar operators was found to introduce significant questions of interpretation for the data. Chapter IX presents the observed effects of intra-group shifts of performance.

CHAPTER VI

DISCRIMINATION WITH DICHOTOMOUS BANDS OF NOISE

6.1 General

In studies of the intelligibility of speech in noise, it is found that a correspondence can be established between the cutoff frequency and listener performance (Pollack, 1948). Furthermore, there seems to be a definite relationship between performance in various audiometric tests and the trained speech recognition process (Stevens and House, 1972). For this reason, one of the principal areas of investigation centered around the effect of high pass filtering on the classification performance against marine sounds.

The study consisted of four experiments involving recordings of marine sounds. In addition, there were three experiments using customized signal patterns. These latter experiments were designed to test performance for signal patterns which had dichotomous features approximating those in the marine source discrimination task.

This chapter expands on the model of the feature extraction and hypothesis testing presented in Chapter II. The detectability of the dichotomous feature is derived and results are compared to experiments conducted by others using signal detectability.

6.2 A Model for Noise Bands Dichotomous

When two pattern classes differ in that one $(\omega_{.H..})$ has the low frequencies eliminated by filtering, the dichotomous feature will be defined to be $\Omega_{LP..}$, a feature which embodies all of the low frequency information. For realistic signals, this feature is associated with a number of source auditory parameters. Typically, the parameters include:

a. A band limited noise with non-uniform spectral density over a range of frequencies. At the low frequency end, the band is determined by the diminution of auditory acuity at near infra-sonic frequencies or by the reproduction medium. At its upper edge, the band is limited by the cutoff frequency of the high pass filter.

b. Temporal variations of amplitude associated with non-stationarity of the noise in the band.

c. More complex temporal changes in the band limited signal such as time-dependent spectral content or tones.

For the purposes of modeling the discrimination process, however, assume that the dichotomous feature $\Omega_{.L..}$ can be characterized strictly as a stationary noise with rectangular bandwidth W_{eff} and corresponding spectral level S_{eff}. The discrimination task then reduces to a hypothesis test for the detection of the dichotomous feature. The two classification decisions are then:

a. Decide $\hat{\omega}_{.H..}$ whenever the null hypothesis is not rejected. Hence, H_0 corresponds to feature $\Omega_{.L..}$ absent.

b. Decide $\hat{\omega}_{.}$ whenever the null hypothesis is rejected. Or, H_1 corresponds to feature $\Omega_{.L..}$ present.

The detectability of a band of noise was discussed in Chapter II. The optimum performance was given by Equation (2.6).

The ability of listeners to detect bands of noise in a noise background was experimentally investigated by Green (1960a). Green found that the observed detectability in this situation is given by:

$$d'_{obs} = a\, d'_{opt} \quad , \quad 0.25 < a < 0.33 \quad ,$$

where a depended on the subject. This finding was seen to hold over a large range of bandwidth and duration of the stimulus. In these experiments, the SNR was varied until the performance, as measured by $P(C)$, attained a desired level. In computing d'_{obs}, he used SNRs which gave a $P(C)$ of about 0.75.

Figure 9 is a compilation of all of Green's findings for the detection of a noise band in noise. The data shown includes data taken over a range of SNRs so as to map out the psychometric function; the behavior of $P(C)$ with stimulus intensity relative background noise. The data was collected using five subjects and with the following set of parameters:

f_c = 400, 800, 1500, 2500, 4500, 6000 Hz

W = 655, 3862, 5143 Hz

T = 3, 10, 30, 100, 300 msec.

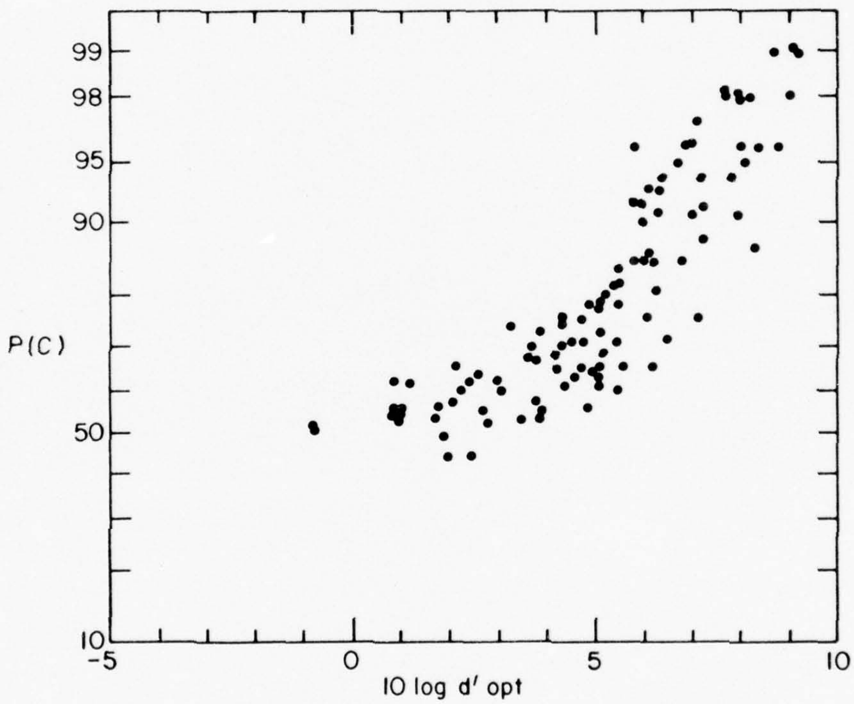


Figure 9. Summary of Green's Results for the Detection of a Noise in Noise.

The parameter f_c is the band center frequency.

These results are for a two alternative, forced choice (2AFC) test and the data is normalized using the equation

$$d'_{opt} = (WT)^{1/2} \frac{\sigma_S{}^2}{\sigma_n{}^2} [1/2(\sigma_S^2/\sigma_n^2)^2 + (\sigma_S^2 + \sigma_n^2) + 1]^{-1/2} . \quad (6.1)$$

Extending these results to the model of discrimination as one of feature extraction and hypothesis tests on the energy in the feature, the model predicts the following:

a. The discrimination performance is predicted to be independent of which band is dichotomous but depends on the effective bandwidth W_{eff}.

b. Given a knowledge of the effective time duration of the stimulus T, the performance should be relatable to that observed by Green.

c. The SNR required to distinguish the signal patterns is further predicted to be unrelated to other parameters of the sound which are common to both sound patterns.

An immediate difficulty in this approach is the definition of the effective time T, and although the SNR was maintained constant for approximately two seconds between step changes, the listener could have integrated over successive steps. Alternately, a time constant considerably shorter than two seconds may apply. While the expression for d'_{opt} implies continuously increasing detectability with time T, it has been found that a listener extracts all useful information from an acoustic stimulus in the first few hundred

milliseconds (Tanner and Sorkin, 1972). Green also observed that

$$d'_{obs} \text{ for } T = 1000 \text{ msec} \simeq d'_{obs} \text{ for } T = 300 \text{ msec.}$$

A detailed analysis of the response behavior for the modified threshold procedure indicates that listeners exhibit comparable behavior in this type of experiment. The number of observed responses as a function of delay from the step change in SNR is plotted in Figure 10. This data is for 160 responses chosen at random from among student events, and for 180 responses from Navy personnel events. In Figure 10, the independent variable τ is the delay to respond from the onset of the step change in SNR. The peak in the number of responses near 600 msec is significant at the 0.1 level for student and sonar operator events under the assumption of a Poisson distribution for the number of responses per 100 msec of delay. This observation tends to support the view that each step change in SNR can be treated as a detection opportunity. After a delay, the listener has extracted all of the additional information provided by the change in detectability associated with the change in SNR and either makes a terminal decision or defers the decision until another such change occurs. It is noteworthy that this result is observed in spite of the fact that listeners were unable reliably to describe the way in which the SNR changed when asked to verbalize the auditory sensation. They were unable consciously to tell if the SNR changed continuously or in steps!

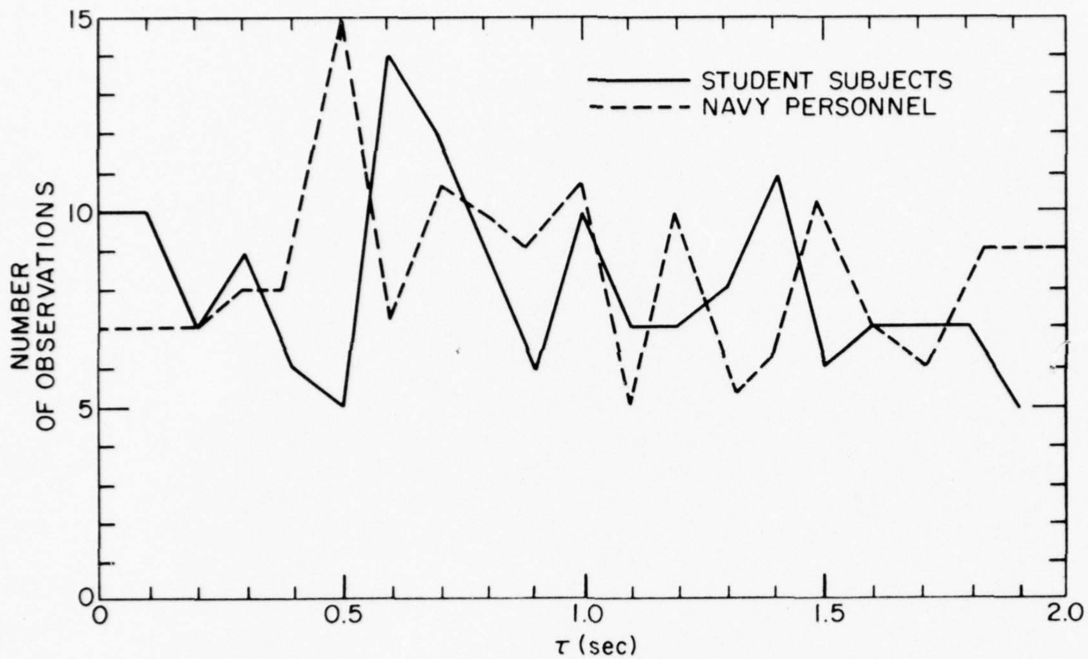


Figure 10. Distribution of Number of Responses as a Function of Elapse Time from a Step Change of Signal-to-Noise Ratio.

On the basis of these data, and in keeping with models of the auditory process, it was inferred that an integration time of 300 to 500 msec is reasonable (Dubrovskii and Tumarkina, 1967; Bauer and Torick, 1966). An additional 100 msec delay due to listener response time would account for the observed peak in subject response delays. The data will subsequently be reduced assuming that the effective time is 500 msec.

When the equivalent square bandwidth W_{eff} is determined for the three marine sounds highpass filtered at 707 Hz, it is found that most of the energy in the dichotomous feature $\Omega_{.L30}$ falls in an octave band centered nominally about 500 Hz. Hence,

$$W_{eff} \cong 172 \text{ Hz for } \Omega_{.L27} .$$

At this point, all of the factors needed to compute d'_{cpt} for the model assumed are available provided that the SNR to respond is known. However, Stallard and Leslie (1974) conclude on theoretical grounds that the difference between a 2AFC experiment, as is Green's experiment, and detection performance in a sonar or other real-world situation should be about 5.4 dB. Their reasoning is as follows:

a. The effect of frequency uncertainty because two bands of noise must be attended to introduces some decrease in detectability for the band of noise to be detected. The other band of noise which must be attended to is one used as a reference to judge the background level.

b. Time uncertainty about the onset time of the signal also has the same effect. This uncertainty is due to the fact that such real-world tasks do not have the time periods during which the signal is present precisely delineated as is the case in 2AFC tests.

c. The fact that the task in the modified threshold situation is one of deciding "yes" the feature is present, or "no" the feature is not present adds an additional factor of $\sqrt{2}$ to the detectability equation. This difference in efficiency of YN and 2AFC tests is equivalent to 1.5 dB in the log detectability equation.

For these reasons, the detectability observed in the modified threshold tests (d'_{mt}) is expected to be:

$$d'_{mt} \cong 3.5\, d'_{obs} \tag{6.2}$$

theoretically, where d'_{obs} is used to identify the results observed by Green. Or,

$$10 \log d'_{mt} \cong 10 \log d'_{obs} + 5.4 \text{ dB} \quad . \tag{6.3}$$

This theoretical performance is predicted if, for each value of SNR, the listener uses only the information available in the auditory stimulus at that time. However, if it were possible for the listener to integrate over successive observations as the SNR increases, this would be useful information and the uncertainty would be reduced, or the detectability would be higher (Reisbeck, 1963). In fact, Swets

and Green (Swets, 1964) have been able to demonstrate that, in very specialized circumstances, listeners can integrate information in successive observations. In general, however, they note that: "This analysis leaves little doubt that the assumption of no integration over successive observations is a good one . . . ." (Swets, 1964, p. 239).

Note, however, that the thresholds γ_A and γ_B applicable to sequential hypothesis testing may change with time or may be related to other factors involved in the test. These thresholds affect the probabilities of the various courses of action but not the form of the applicable detectability equation for d'_{opt}.

In summary, applying the model of a discrimination as a hypothesis test about the presence of a dichotomous feature, the observables in an experiment using the modified threshold are $\overline{SNR}$, the mean signal-to-noise ratio to make a terminal decision, and $P(C)$, the probability of a correct classification. This probability will be related to the criterion used to make the terminal decision, in this case, "reasonably certain."

The probability $P(C)$ is given by

$$P(C) = P(\hat{\omega}_{.}|H_1)\ P(H_1) + P(\hat{\omega}_{.H..}|H_0)\ P(H_0) \tag{6.4}$$

since a correct decision consists of deciding correctly that the feature $\Omega_{.L..}$ is present (H_1) or that the feature is absent (H_0). The conditions H_0 and H_1 are of course mutually exclusive and exhaustive.

The above probability $P(\hat{\omega}_{.}|H_1)$ corresponds to the "hit rate" in event-response formulations leading to the definition of receiver operating characteristics (Egan, 1975). The rate of incorrect acceptance or "false alarm" rate is then

$$P(FA) = P(\hat{\omega}_{.}|H_0) \quad .$$

Furthermore, referring to Figure 2b, the error regions under the two distributions are:

$$\text{a.} \quad P(FA) = P(\hat{\omega}_{.}|H_0) = \alpha_t = \int_{\gamma_A}^{\infty} p_{X|H_0}(X|H_0)\,dx \tag{6.5}$$

$$\text{b.} \quad P(\hat{\omega}_{.L..}|H_1) = \beta_t = \int_{0}^{\gamma_B} p_{X|H_1}(X|H_1)\,dx \quad . \tag{6.6}$$

6.3 Experimental Results for High Pass Filtered Marine Sounds

The four experiments using high pass filtering of marine sounds are summarized in Table 2. The treatments investigated are further detailed in Table 5 where the number of events per treatment is also shown. Events for which the subjects did not respond have been omitted from Table 5.

From Table 5, it is observed that some treatments differ only in the order in which the signal patterns are presented during the exposure sequence. For example, Treatments I.1 and I.2 both have the unfiltered marine sound (ω_1) as the probe signal but differ in the fact that, for Treatment I.1, this signal pattern is presented first and is identified as signal A to the subjects. Also, some of

TABLE 5

SUMMARY OF TREATMENTS FOR STUDIES OF THE
EFFECT OF HIGH PASS FILTERING OF A MARINE SOUND

Exposure Sequence A	Exposure Sequence B	Probe	Treatment No./No. of Events: Student Subject	Sonar Operators: Inexperienced	Sonar Operators: Experienced
ω_1	ω_{1H30}	ω_1	I.1/11	-	
ω_{1H30}	ω_1	ω_1	I.2/21	I.5/8	
ω_1	ω_{1H30}	ω_{1H30}	I.3/24	-	
ω_{1H30}	ω_1	ω_{1H30}	I.4/9	I.6/9	
ω_2	ω_{2H30}	ω_2	II.1/9	II.5/8	
ω_{2H30}	ω_2	ω_2	II.2/35	-	
ω_2	ω_{2H30}	ω_{2H30}	II.3/9	II.6/4	
ω_{2H30}	ω_2	ω_{2H30}	II.4/9	-	
ω_2	ω_{2H27}	ω_2	-	IV.1/12	
ω_{2H27}	ω_2	ω_2	-	-	
ω_2	ω_{2H27}	ω_{2H27}	-	IV.2/6	
ω_{2H27}	ω_2	ω_{2H27}	-	-	
ω_3	ω_{3H30}	ω_3	III.1/33	III.5/15	III.9/17
ω_{3H30}	ω_3	ω_3	III.2/9	III.6/10	III.10/9
ω_3	ω_{3H30}	ω_{3H30}	III.3/27	III.7/15	III.11/14
ω_{3H30}	ω_3	ω_{3H30}	III.4/32	III.8/4	III.12/6

the treatments have only a few events because results had to be discarded due to equipment malfunction or noise at the test site. Some treatments which were planned to be tested in greater detail with sonar operators were not done because of operators not coming on time or failing to show up altogether. Treatments II.6, III.8, III.12, and IV.2 probably have too few data points to allow even testing differences of means on SNR.

The first order statistics for Experiment I are tabulated in Table 6. This table shows the mean signal-to-noise ratio ($\overline{SNR}$) to respond, and the sample standard deviation S_{SNR}. Likewise, these same data are contained in Table 7 for Experiment II and in Table 8 for Experiment III. Only this latter experiment has enough data to allow comparisons between the two groups of sonar operators, those with less than two years experience and those with two years experience or more.

It is seen in all student cases that treatments which differed only in the order of the exposure signals, and which had the dichotomous feature present during the response period, yielded comparable values of $\overline{SNR}$ and could therefore be pooled in the data analysis and treated as equivalent. However, in every student experiment of this kind, there was a statistically significant difference in results for the treatments where the probe was missing the dichotomous feature ($\omega_{.H30}$) and the order of signal exposure was different. This difference yielded significant t scores which were:

TABLE 6

FIRST ORDER STATISTICS FOR EXPERIMENT I,
A DISCRIMINATION TEST WITH BAND OF NOISE DICHOTOMOUS

Treatment	$\overline{SNR}$	S_{SNR}	N
I.1	1.37	2.72	11
I.2	1.80	2.65	21
I.3	0.71	2.51	24
I.4	-2.72	2.74	9
I.5	5.10	2.64	8
I.6	4.99	3.03	9

TABLE 7

FIRST ORDER STATISTICS FOR EXPERIMENT II,
SIGNAL PATTERNS ω_2, ω_{2H30}

Treatment	$\overline{SNR}$	S_{SNR}	N
II.1	-0.42	2.86	9
II.2	0.77	3.83	35
II.3	1.71	2.02	9
II.4	-2.00	3.37	9
II.5	7.19	3.53	8
II.6*	5.20	2.64	4

*Of doubtful significance

TABLE 8

FIRST ORDER STATISTICS FOR EXPERIMENT III

Treatment	$\overline{SNR}$	S_{SNR}	N
III.1	5.58	2.57	33
III.2	4.73	3.17	9
III.3	4.09	3.00	27
III.4	2.45	3.39	32
III.5	7.28	3.53	18
III.6	3.56	6.87	7
III.7	7.66	4.31	16
III.8*	4.25	5.15	4
III.9	10.63	3.42	17
III.10	6.37	2.68	9
III.11	10.94	3.26	14
III.12*	5.80	2.35	6

*Of doubtful significance

a. $T = 3.41$ with 31 degrees of freedom (ν) in Experiment I

b. $T = 2.83$, $\nu = 16$ in Experiment II

c. $T = 1.95$, $\nu = 57$ in Experiment III

where

$$T = \frac{(\overline{SNR}_1 - \overline{SNR}_2)}{S_{SNR} \sqrt{(1/N_1) + (1/N_2)}} \quad . \tag{6.7}$$

In comparisons such as these, the T score is significant at the 0.1 level whenever

$$T > t\,(0.95, \nu)$$

and $t\,(0.95, \nu)$ is the single tail t statistic at the 95 percent level with ν degrees of freedom (Weinberg and Schumaker, 1962).

This effect is difficult to explain since it goes counter to expectations based on "recency" of the training stimulus (Klatzky, 1975). Both groups of sonar operators also show a similar dependence on order of exposures. Pooling of data without regard to exposure order can therefore lead to a type I error when testing if there is a significant difference in $\overline{SNR}$ to respond when the signal patterns $\omega_{.}$ and $\omega_{.H..}$ serve as the probe. Type I errors are committed whenever the null hypothesis is rejected when it is true.

Referring again to Table 5, it is seen that Treatments I.1 and I.2, I.3, and I.4, etc., are similar in exposure order and only differ in the probe used. Using the case where $\omega_{.}$ is exposed first in the treatment, the point estimate of $\overline{SNR}$ is obtained under hopefully

similar experimental conditions. The results are plotted in Figures 11, 12, 13, and 14 for both student subjects and sonar operators. In some cases, however, the case where signal pattern $\omega_{H..}$ was exposed first is used instead. This is the case when comparison to limited sonar operator data was desired as is the case for Experiment I. These figures show the signal excess over noise in dB in third-octaves for the total frequency range. Note that a change in $\overline{SNR}$ in the modified threshold experiment leads to a simple translation of the signal excess curve up or down. The figures graphically demonstrate how far "out of the noise" the signal must be for the listeners to obtain enough data to commit to a terminal decision.

In addition, it is seen that in most cases, the $\overline{SNR}$ to discriminate when the dichotomous feature was present is comparable to that needed to discriminate when the dichotomous feature was absent. In the cases where the $\overline{SNR}$ was significantly different for the two probe conditions, the observation may be the result of the major effect of exposure order.

For those treatments where there was a significant difference between discrimination for the two probe conditions, the data indicate that the $\overline{SNR}$ for the terminal decision was <u>lower</u> when the dichotomous feature was absent. These treatments were:

a. I.2 and I.4, $T = 4.24$ for 28 degrees of freedom (ν)

b. II.2 and II.4, $T = 1.98$, $\nu = 42$

c. III.1 and III.3, $T = 2.09$, $\nu = 59$.

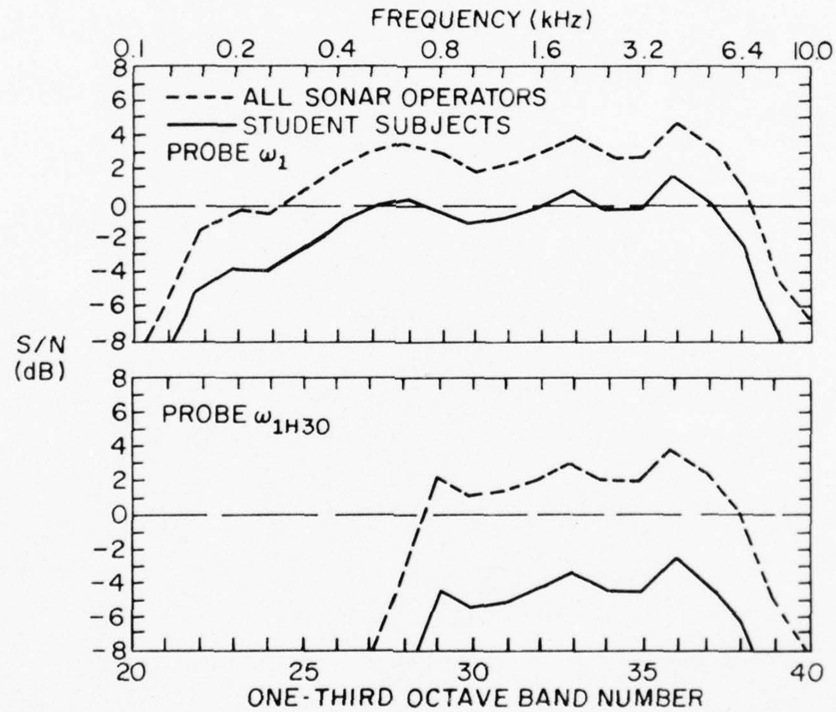


Figure 11. Signal Excess in One-Third Octave Bands at the Terminal Decision, Signal Pattern ω_1, ω_{1H30}.

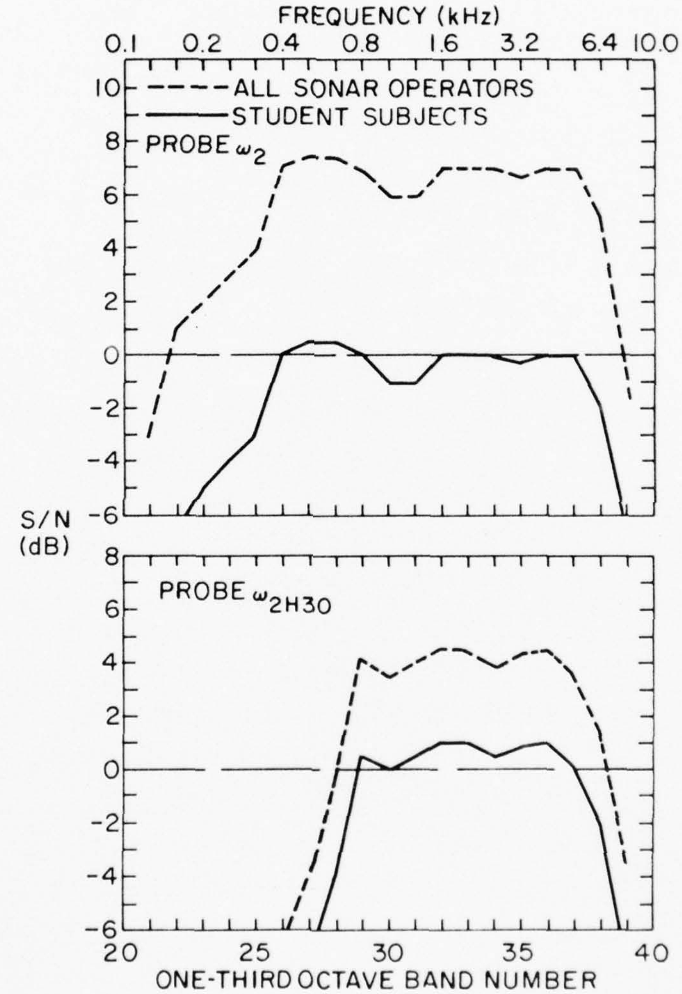


Figure 12. Signal Excess at the Terminal Decision for Signal Patterns ω_2 and ω_{2H30}.

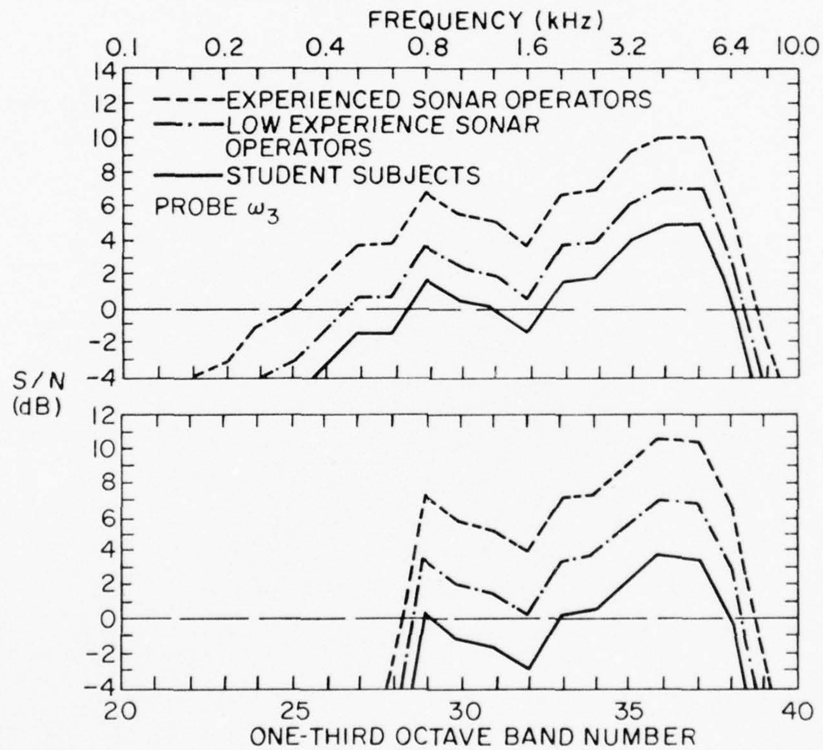


Figure 13. Signal Excess at the Terminal Decision for Signal Patterns ω_3 and ω_{3H30}.

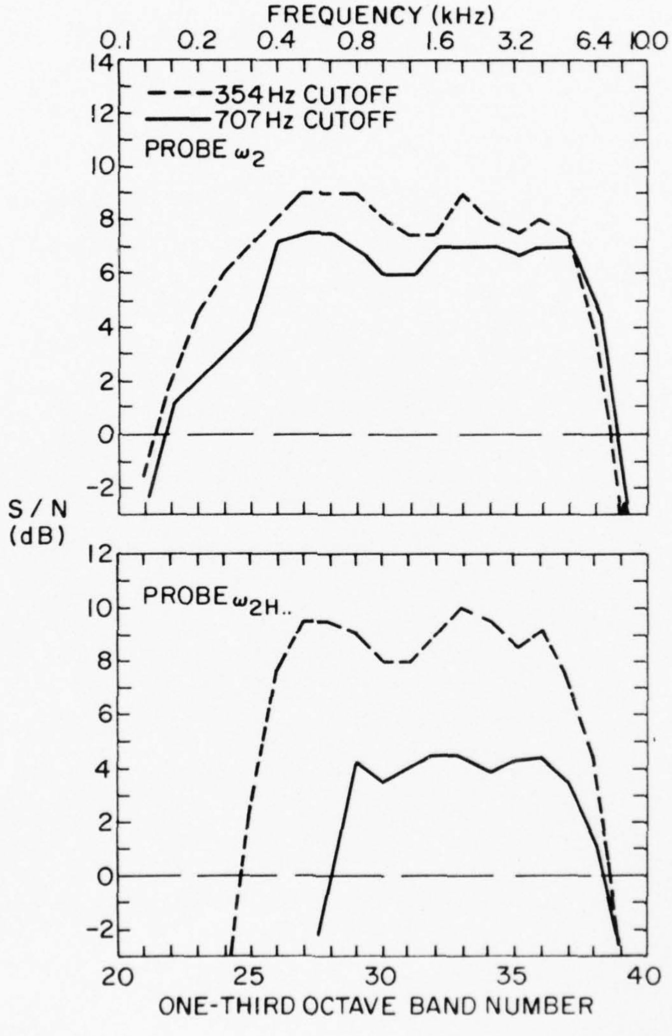


Figure 14. Signal Excess at the Terminal Decision for Signal Pattern ω_2 Showing the Effect of Varying the Cutoff Frequency.

The data in Table 8 also show that the population of listeners consisting of sonar operators with little at-sea experience exhibit a dramatically higher standard deviation in SNR to respond than do either University students or highly experienced operators. To test if this standard deviation (or variance) is in fact significantly greater, the F statistic is computed (Walpole and Myers, 1972). However, unlike the T statistic, which is very distribution insensitive, the F statistic is valid only whenever the distribution of observations can be assumed normal. The observed SNR to respond is plotted in Figure 15 for sonar operators with Experiment III. The figure shows that an assumption of normally distributed observations is not unwarranted. The observed difference in the variance exhibited by the low experience sonar operators against that of the University students and experienced sonar operators was in no case significant at the 0.1 level. Hence, the difference could simply be due to chance.

The other quantity measurable in the modified threshold experiments is the probability of various decision alternatives. The experimental results are summarized in Table 9. This table gives both the point estimate and 90 percent confidence interval on the probabilities.

The two alternatives corresponding to correct decisions are:

a. $P(\hat{\omega}_{.}|H_1)$, the probability of correct acceptance, or hit rate.

b. $P(\hat{\omega}_{.H..}|H_0)$, the probability of correct rejection when the dichotomous feature $\Omega_{.L..}$ is absent.

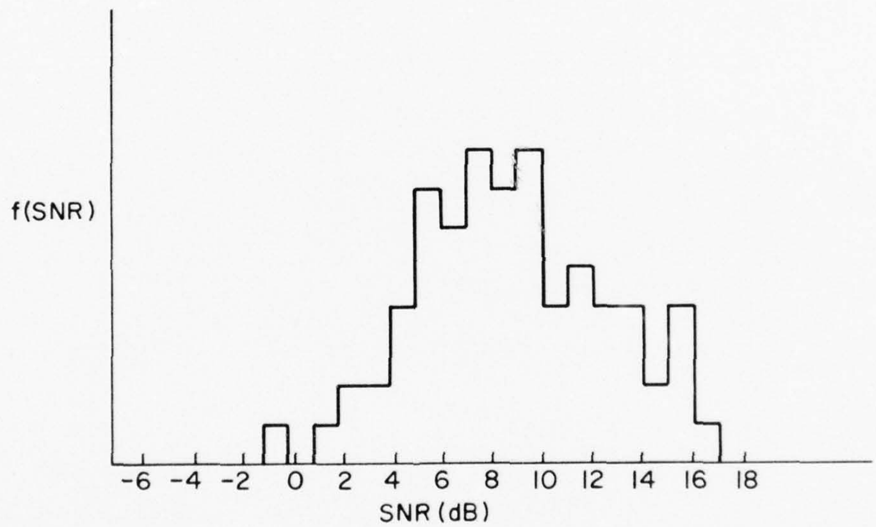


Figure 15. Distribution of Signal-to-Noise Ratio to Respond for Navy Personnel, Experiment III.

TABLE 9

SUMMARY OF PROBABILITIES EXPERIMENTALLY MEASURED
FOR DISCRIMINATION WITH DICHOTOMOUS BANDS OF NOISE

Experiment	Subject Group	Point and Interval Estimates		
		$P(\hat{\omega}_{.} \mid H_1)$	$P(\hat{\omega}_{.H..} \mid H_0)$	P(C)
I	Students	0.69, 0.54 - 0.82	0.79, 0.66 - 0.89	0.74, 0.64 - 0.82
II	Students	0.75, 0.56 - 0.80	0.61, 0.41 - 0.79	0.71, 0.62 - 0.80
III	Students	0.48, 0.32 - 0.65	0.71, 0.61 - 0.81	0.63, 0.55 - 0.71
	Inex. Sonar Operators	0.54, 0.37 - 0.70	0.55, 0.36 - 0.73	0.51, 0.39 - 0.64
	Exp. Sonar Operators	0.54, 0.37 - 0.70	0.74, 0.55 - 0.89	0.62, 0.50 - 0.74
VIII	All Sonar Operators	0.77, 0.63 - 0.88	0.63, 0.45 - 0.78	0.70, 0.59 - 0.80
IX	All Sonar Operators	0.68, 0.52 - 0.82	0.82, 0.64 - 0.95	0.73, 0.62 - 0.84

In addition, the resulting probability of a correct response, correct acceptance or rejection, is also shown. The table clearly shows how large the confidence intervals on probabilities is unless a large number of data points is available. Only treatment results with sufficient data points to give some credibility to the point estimate are shown. Data are pooled across presentation order.

6.4 Relationship Between Results for High Pass Filtering and the Model

To determine how well the experimental results relate to the model, the data are compared to those obtained by Green (1960a) as described in Section 6.2. In that section, the normalization equation was presented and the effective bandwidth W_{eff} was shown. The effective spectral level S_{eff} is then obtained by noting the $\overline{SNR}$ to respond experimentally. Once this value is known, the spectral level of the dichotomous band of noise $\Omega_{.L..}$ is obtained directly from the calibration data.

The point and interval estimates for the effective ratios of signal-to-noise power in the dichotomous band

$$S^2_{eff}/N^2 = \sigma^2_{\Omega}/\sigma^2_n$$

are shown in Table 10. The table also shows the point and interval estimates on the quantity

$$10 \log d'_{mt} \quad ,$$

TABLE 10

RESULTS OF MODIFIED THRESHOLD PROCEDURE FOR CLASSIFICATION UNDER DICHOTOMOUS FEATURES, HIGH PASS CASE

Feature	Subjects	10 log $(\sigma^2_\Omega/\sigma^2_n)$	Point and Interval Estimate, 10 log d'_{mt}
Ω_{1L30}	Students	1.0, 0.5/1.5	9.1, 8.8 - 9.5
	All Sonar Operators	5.2, 4.3/6.1	11.2, 10.9 - 11.5
Ω_{2L30}	Students	.7, .2/1.3	8.9, 8.6 - 9.3
	All Sonar Operators	6.7, 5.5/7.9	11.6, 11.3 - 11.9
Ω_{3L30}	Students	0.6, 0.2/1.0	8.9, 8.6 - 9.1
	Inex. Sonar Operators	1.9, 1.2/2.6	9.7, 9.3 - 10.1
	Exper. Sonar Operators	5.2, 4.6/5.8	11.2, 11.0 - 11.4
Ω_{2L27}	All Sonar Operators	4.3, 3.3/5.3	9.4, 9.0 - 9.8

the observed detectability for the feature normalized using Equation (6.1). The subscript mt is used to designate data obtained using modified threshold techniques.

Note that the dichotomous feature Ω_{2L27} resulting from high pass filtering at 354 Hz exhibits a detectability which is comparable to that found for the features $\Omega_{\cdot L30}$ resulting from high pass filtering signal patterns ω_1, ω_2, and ω_3 at 707 Hz.

When all of the data for high pass filtering of marine sounds are compared, it is seen that the performance for the various signal patterns is comparable. The results are plotted in Figure 16 along with a fourth order polynomial regression fit to Green's data (Green, 1960a; McGill, 1968). The theoretically predicted correction for a sonar or other industrial environment detection problem is also shown (Stallard and Leslie, 1974).

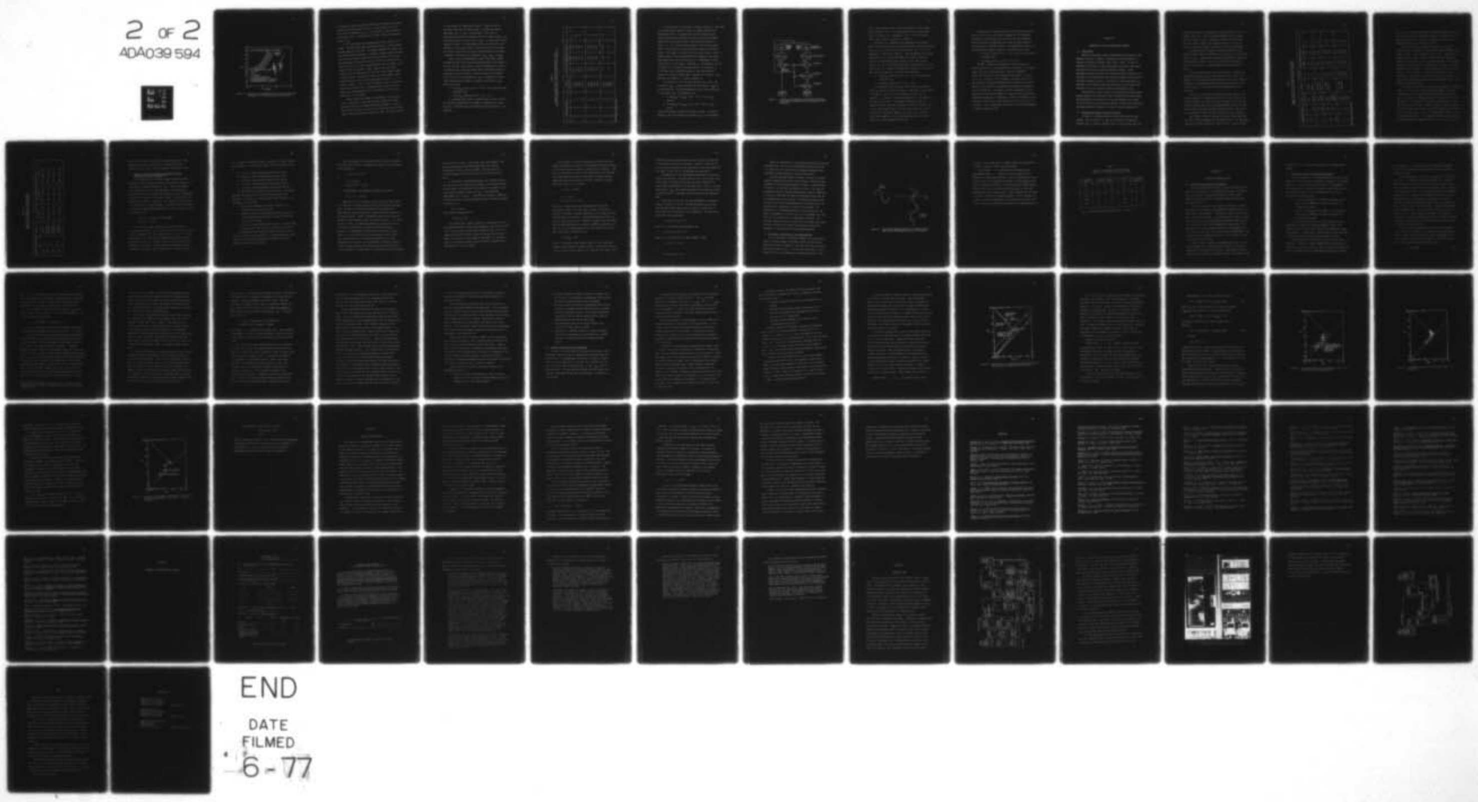
Figure 16 shows that University student listeners' performance falls very near that predicted by the model. Also, sonar operators seem to consistently require a higher $\overline{SNR}$ to respond but they are not able to improve their performance in terms of P(C) proportionately. In the one experiment where data of inexperienced vs experienced sonar operators were sufficient, the experienced operators did show an apparent improvement in P(C) at a higher $\overline{SNR}$. This difference in P(C) is not statistically significant, however.

The consistency of these findings, in spite of subject differences, is gratifying especially when it is recalled how many simplifying assumptions are embodied in the model of the listener

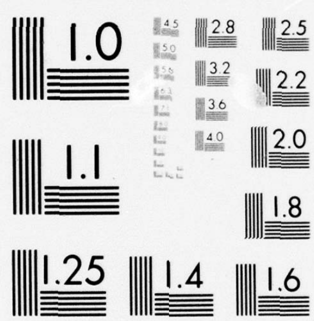
AD-A039 594 PENNSYLVANIA STATE UNIV UNIVERSITY PARK APPLIED RESE--ETC F/G 6/16
AN EXPERIMENTAL TREATMENT OF AUDITORY DISCRIMINATION OF COMPLEX--ETC(U)
NOV 76 C P JANOTA N00017-73-C-1418

UNCLASSIFIED TM-76-301 NL

2 OF 2
ADA039 594



END
DATE
FILMED
6-77


1.0
4.5
5.0
5.6
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

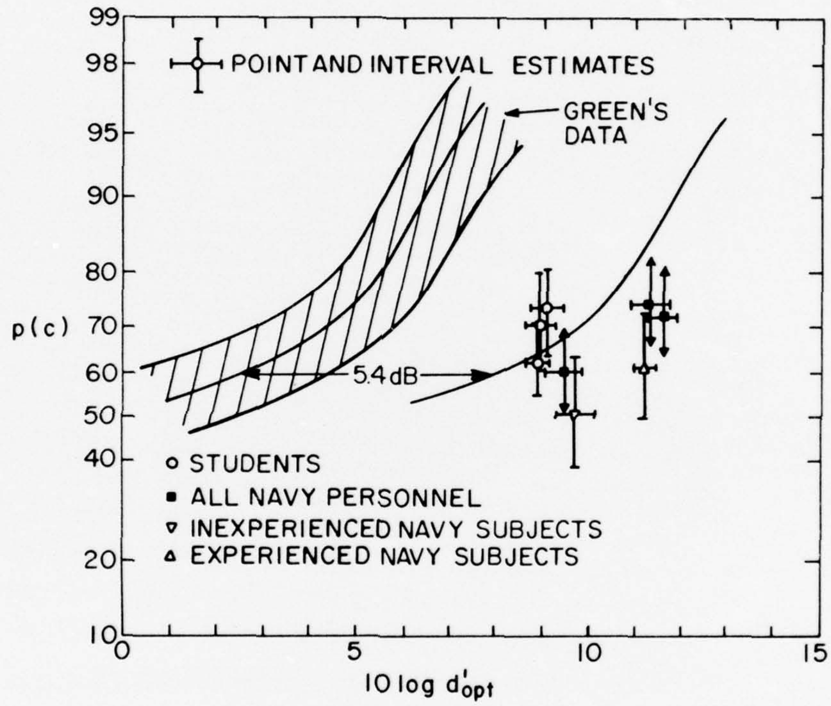


Figure 16. Comparison of Experimental Results with the Predicted Performance for the Detection of Noise in Noise.

as performing a feature extraction followed by a hypothesis test based on energy in the dichotomous feature. Just the uncertainty of the exact effective bandwidth W_{eff} to use, and that of the uncertainty of effective time T, could account for nearly 2 dB in the resulting value of d'_{mt}.

So far, the problem has been reducible to one of detection of a feature. Certainly the present discrimination task is a more complex one than implied in the model. For example, whenever all energy above the cutoff frequency used here is removed, the model predicts the same results since it does not make use of the information in this band. The performance is derated by a factor to allow for listening to two bands of energy but no other assumptions about how this energy in the fixed portion of the signal is used. However, were the higher frequency features removed in practice, the listener behavior would be much different. The present results show that the $\overline{SNR}$ to make a terminal decision is not much different when the feature $\Omega_{.L..}$ is present or absent. With no other information, however, the listener would never be able to make a terminal decision with the feature absent since nothing would serve as a reference as to how far the signal was out of the noise.

Clearly, then, the listener is simultaneously using various types of information in the signal pattern to make the terminal decision required in the modified threshold paradigm. In order to obtain better insight as to the role of the various features of the signal, laboratory generated signals were used. The characteristics

of these signals are summarized in Table 3. Table 4 lists the experiments performed with these sounds. Of these experiments, Experiments VIII, IX, and X are applicable to the study of discrimination where the dichotomous feature is a band of noise. The dichotomous feature in these cases is Ω_{B27}, an octave band of noise centered at 500 Hz. The fixed feature was an octave band of noise centered at 4 kHz and was chosen to have three distinct variations which serve as the factors for which the discrimination performance is measured. This feature is designated Ω_{7B36}, Ω_{7B36}, or Ω_{9B36}.

Table 11 lists the experimental results using these signals. The exposure order effect previously mentioned seems to be present in this data as well. Note, however, that for Experiment VIII there is a significant and consistently higher SNR to discriminate when the dichotomous feature was absent. Using the results for Treatments VIII-1 and VIII-2, the detectability d'_{mt} is obtained for the two cases differing in the absence or presence of the feature Ω_{B27}. These results were:

a. 10 log d'_{opt} under H_1 is $9.4 < 10.0 < 10.4$ at the 90 percent confidence level

b. 10 log d'_{opt} under H_0 is $10.8 < 11.1 < 11.4$.

To investigate this effect in greater detail, the way in which the listener recognizes that the signal is sufficiently far above the noise to determine that the dichotomous feature is absent is now addressed.

TABLE 11

SUMMARY OF RESULTS FOR STUDIES OF DISCRIMINATION USING
LABORATORY GENERATED SOUNDS, AN OCTAVE BAND OF NOISE DICHOTOMOUS

Exposure Sequence		Probe	Treatment	Experimental Results		N
A	B	$\omega_{.}/H_0$ or H_1	No.	$\overline{SNR}$	S_{SNR}	
ω_5	ω_7	ω_5/H_1	IIX.1	0.69	3.52	14
ω_7	ω_5	ω_5/H_1	IIX.2	-2.30	3.48	16
ω_5	ω_7	ω_7/H_0	IIX.3	3.32	3.53	21
ω_7	ω_5	ω_7/H_0	IIX.4*	0.93	2.85	3
ω_6	ω_7	ω_6/H_1	IX.1	6.39	5.06	8
ω_7	ω_6	ω_6/H_1	IX.2	7.36	2.62	20
ω_6	ω_7	ω_7/H_0	IX.3*	4.03	2.70	3
ω_7	ω_6	ω_7/H_0	IX.4	7.84	3.23	14
ω_8	ω_9	ω_8/H_1	X.1	6.18	3.26	15
ω_9	ω_8	ω_8/H_1				-
ω_8	ω_9	ω_9/H_0	X.2	5.29	3.77	7
ω_9	ω_8	ω_9/H_0				-

*Of doubtful significance

A possible model of the process is shown in Figure 17. This model is an application of Neisser's theory for pattern recognition by humans (Reed, 1973). The model shows that the listener must do feature extraction and hypothesis testing on both the dichotomous feature and the invariant feature to decide that the feature is absent, i.e., H_0. The feature extraction is here shown as being done in parallel which is one of the salient features of Neisser's theory. This aspect of the model is not critical, however, and it is intended only to show the possible decision sequence. This simple model demonstrates that the decision H_1 is independent of other characteristics of the signal. Threshold γ_B corresponds to the one in Chapter II for sequential decision making. The threshold γ_C is the same kind of threshold applied to a decision about the presence of the high frequency band. The decision constants γ_D, γ_E determine the threshold γ_A shown in Chapter II.

From Figure 17, it is seen that the decision to report that the feature Ω_{B27} is absent requires also that feature Ω_{7B36} be at least detected. From the experimental results, the detectability of this feature was, for Experiment VIII,

a. $10 \log d'_{opt}$ for $\Omega_{7B36} = 10.3 < 11.0 < 11.6$ for Ω_{B27} present,

b. $10 \log d'_{opt}$ for $\Omega_{7B36} = 12.5 < 13.1 < 13.6$ for Ω_{B27} absent

where the 90 percent confidence intervals are shown. It appears, therefore, that the threshold for deciding H_0 on the observables

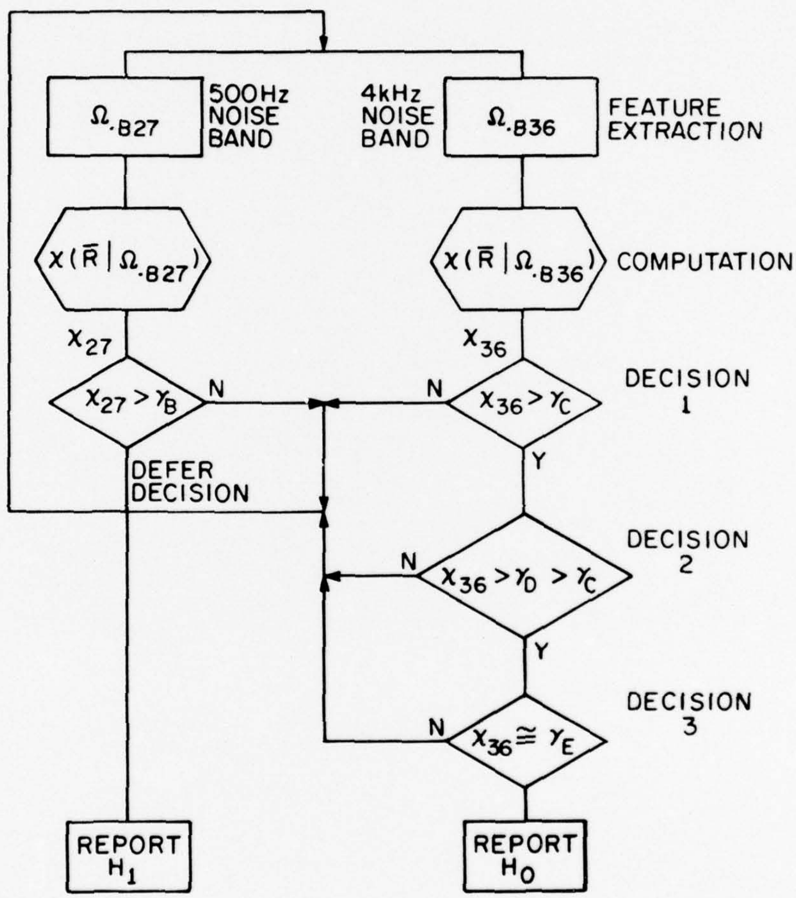


Figure 17. A Model of Discrimination Incorporating Additional Clues Needed to Decide that the Dichotomous Feature is Absent.

from the high frequency band is about 3 dB higher than that needed to detect that band. Here it is assumed that the high frequency band requires a detectability comparable to that for the low frequency band to satisfy the first decision in Figure 17.

Experiment IX is identical to the experiment detailed above except that the spectral level of the high frequency band was increased 6 dB relative to that in the low frequency band. The data for this experiment is summarized in Table 11. Since the $\overline{SNR}$ measured in these experiments is for a third-octave band at the higher frequencies relative to the background noise, the data in the table actually relate to how far out of the noise the high frequency band was at the terminal decision. Because of limited data, Treatments IX.2 and IX.4 were used in the analyses to give the following results:

a. 10 log d'_{opt} under H_1 is $9.1 < 9.4 < 9.8$ at the 90 percent confidence level,

b. 10 log d'_{opt} under H_0 is $9.2 < 9.7 < 10.2$.

The experimental data show that in this case the performance under H_1 and H_0 are nearly identical. This tends to lend credence to the model of the process as postulated. For this case, the threshold γ_D (Figure 17) would be surpassed by the observables in the high frequency band. The listener must make an additional comparison shown as Decision 3 in Figure 17 which is related to how far the high frequency band is out of the noise and which seems to require an accurate mental image of the sound. The listener in effect must remember how the low and high bands were related and match this to the background noise as well.

In Experiment X, the fixed feature was an amplitude modulated band of noise which was very easily recognized in the background. According to the model discussed in this section, this experiment should be analogous to Experiment IX in that the feature common to both signals has a high detectability. The experimental results for the detectability needed to commit to a terminal decision were:

a. $10 \log d'_{opt}$ under H_1 is $11.1 < 11.4 < 11.6$ at the 90 percent confidence level,

b. $10 \log d'_{opt}$ under H_0 is $10.5 < 11.1 < 11.5$.

The $\overline{SNR}$ to respond was comparable for the cases where the dichotomous feature was absent and when it was present. This finding agrees with what was predicted above. However, the $\overline{SNR}$ was significantly higher than that observed in Experiments VIII and IX for the H_1 situation. One is tempted to attribute this effect to the presence of the highly detectable feature interacting with the ability of the listener to concentrate his attention on the band of noise to be detected. The subjects, in fact, complained that they were having a hard time attending to the task because the amplitude modulated component was so "monotonous." However, the amount of data available as the result of these studies does not allow for any such conclusion. This is especially true because of the seeming dependence of results on exposure order.

CHAPTER VII

EXPERIMENTS WITH OTHER DICHOTOMOUS FEATURES

7.1 Introduction

Studies investigating speech intelligibility have shown that the temporal structure of signals is important in the determination of listener performance (Licklider, 1951; Miller, 1948). In order to investigate this aspect of noise-like sounds, two sets of experiments were performed using the modified threshold procedure. These experiments compared classification performance when the exposure set consisted of two signals which were spectrally similar but differed in the signal envelope structure. The results of these experiments are presented in this chapter along with an analysis of the results from the point of view of a feature extraction and hypothesis test.

Also discussed in this chapter is an experiment using University student subjects only which used "soft" limiting of noise-like sounds. The finding that such distortion of the time waveform of the signal has little effect on its subjective character is noted and similarities to speech distortion tolerance is cited as a possible explanation.

7.2 Changing the Temporal Structure of Sounds

In Table 1 are characterized the marine sounds used in these studies. The two sounds, ω_1 and ω_2, which have a pronounced rhythmic quality, exhibit a temporal shift in the amplitude with time.

This amplitude change, or modulation, is quasi-periodic with a rate of between 7 and 10 Hz. Since the temporal amplitude shift of speech signals is known to convey a significant amount of the information about the signal, changes in the temporal structure of the marine sounds is expected to change its subjective character considerably.

Two experiments were conducted using both student and sonar operator subjects where the modulation of the sound was the dichotomous feature ($\Omega_{.AM}$). An additional experiment using laboratory generated sounds was also included. In this case, the amplitude modulation was a 10 Hz square wave. These experiments are also tabulated in Table 2.

The tape recorded marine sounds were treated to eliminate the amplitude modulation by using a fast automatic gain control. This method, without changing the overall spectrum of the sound, effectively eliminates all amplitude modulations slower than about 70 Hz. The rhythmic quality of the sound is effectively eliminated by this treatment.

The results of modified threshold tests using student or sonar operator listeners are summarized in Table 12. Sonar operators are treated together even though the results in Chapter VI indicate that there may be some significant differences between operators with less than two years experience and highly experienced ones. There is not sufficient data to allow splitting the data by operator sub-group.

The results in Table 12 show that this type of experiment does not seem to exhibit dependence on the exposure order as did experiments where the dichotomous feature was a band of noise. The one instance

TABLE 12

SUMMARY OF RESULTS FOR STUDIES OF DISCRIMINATION
WHERE SIGNALS DIFFER IN THEIR TEMPORAL STRUCTURE

Exposure Sequence		Probe	Subjects	Treatment No.	Experimental Results		N
A	B	$\omega_{.}/H_0$ or H_1			$\overline{SNR}$	S_{SNR}	
ω_{1SN}	ω_1	ω_1/H_1	Student	V.1	-4.13	2.58	10
ω_1	ω_{1SN}	ω_{1SN}/H_0	Subject	V.2	-2.99	3.46	7
ω_1	ω_{1SN}	ω_1/H_1	All Sonar	V.3	2.87	3.30	14
ω_1	ω_{1SN}	ω_{1SN}/H_0	Operators	V.4	3.98	4.56	14
ω_2	ω_{2SN}	ω_2/H_1	All Sonar	V1.1	2.74	2.25	12
ω_{2SN}	ω_2	ω_2/H_1	Operators	V1.2	3.26	4.50	14
ω_2	ω_{2SN}	ω_{2SN}/H_0		V1.3	1.67	2.27	4*
ω_{2SN}	ω_2	ω_{2SN}/H_0		V1.4	6.97	3.33	16
ω_4	ω_5	ω_4/H_1	All Sonar	V11.1	2.34	1.81	8
ω_5	ω_4	ω_4/H_1	Operators	V11.2	1.75	3.50	19
ω_4	ω_5	ω_5/H_0		V11.3	6.67	3.30	15
ω_5	ω_4	ω_5/H_0		V11.4	6.63	2.11	10

*Of Doubtful Significance

where such an effect may be present is between Treatments VI.3 and VI.4 but this is of doubtful significance because of the small number of data points associated with the former treatment and because of one doubtful data point at an unusually low SNR.

Also, as with results presented in Chapter VI, the sonar operator subjects responded at a significantly higher $\overline{SNR}$ than did the student listeners. This difference is about 7 dB. Subsequent tests using a graduate student to repeat the tests show that the results obtained by the student subjects are again attainable. The difference between the two populations of listeners seems to be one of criterion. The sonar operators appear to be more conservative in their terminal decision. Such differences in criterion should reflect in the performance as measured by the probability of a correct decision. In Table 13, the significant probabilities are summarized for tests with temporal features dichotomous.

The data show that the sonar operators' performance, if anything, tends to be somewhat less reliable than the students'. The large confidence intervals preclude disproving the null hypothesis, i.e., there is no difference between populations, at the 0.1 level, however. The value of P(C) exhibited by the students in Experiment V is high and the sonar operator data is somewhat tainted by the fact that some of the less experienced operators exhibited a surprisingly high error rate on this experiment. There is not enough data, however, to allow for definite conclusions. It is noteworthy, however, that two of the most experienced sonar operators made no errors on any of the events

TABLE 13

PROBABILITY EXPRESSIONS FOR EXPERIMENTS
WHERE TEMPORAL FEATURES ARE DICHOTOMOUS

Feature	Experiment	Subjects	Point and Interval Estimate of Probability		
			$P(\hat{\omega}_{.}\|H_1)$	$P(\hat{\omega}_{.SN}\|H_0)$	$P(C)$
Ω_{1AM}	V	Students	0.69, 0.90, 1.00	0.66, 0.83, 0.95	0.73, 0.86, 0.95
Ω_{1AM}	V	All Sonar Operators	0.34, 0.60, 0.84	0.60, 0.80, 0.94	0.55, 0.70, 0.83
Ω_{2AM}	VI	All Sonar Operators	0.76, 0.88, 0.97	0.66, 0.82, 0.93	0.76, 0.85, 0.93
Ω_{4AM}	VII	All Sonar Operators	0.87, 0.96, 1.00	0.87, 0.96, 1.00	0.90, 0.96, 0.99

where the dichotomous feature was the amplitude modulation of the signal. These same operators were very conservative in their responses, however, and routinely responded at high values of SNR.

7.3 Analysis of Results Obtained in Experiments with the Temporal Characteristic Modified

In the experiments with signals treated to remove amplitude modulation vs the recorded marine sound, the dichotomous feature was the presence of amplitude modulation (Ω_{AM}). This feature appears as a repeated burst of noise impressed upon a steady background. Miller and Taylor (Miller, 1948; Miller and Taylor, 1948) investigated the subjective character of this type of signal. They found that the differential threshold for intensity $\Delta I/I$ increases as the duration of the added burst of noise decreases. In the limit where the duration of the added increment exceeds about 250 msec, the performance is given by the Weber fraction

$$\Delta I/I \cong 0.1, \; t > 250 \text{ msec (Green, 1960a)} ,$$

$$10 \log \Delta I/I \cong -10 \text{ dB} .$$

For the marine sounds in question, the natural modulation corresponding to feature Ω_{AM} is in the form of short bursts of noise which are repeated more or less periodically at a rate of 7 to 10 Hz. The duration of the noise pulse is between 25 and 50 msec but with a non-rectangular waveform. Figures 11 and 12 show that the effective bandwidth of the signal vs noise background is approximately 6 kHz from 400 Hz to 6400 Hz at the 3 dB down points. The Weber fraction

$\Delta I/I$ is then the incremental change in intensity vs average intensity just perceptible. Moore and Raab (1975) find that the Weber fraction is:

a. -1.4 dB for 3160 Hz bandwidth and duration 10 msec.

b. -0.9 dB for 1000 Hz bandwidth and duration 10 msec.

c. -5.5 dB for 3160 Hz bandwidth and duration 250 msec.

d. -4.9 dB for 1000 Hz bandwidth and duration 250 msec.

e. -7.3 dB for an 18 kHz bandwidth and duration 250 msec.

For Experiment V, the central 12.5 msec duration of the peak in amplitude was measured to be 4 dB relative the average signal amplitude. Using this fact, and the results shown in Table 12, the following can be calculated:

a. The average signal-plus-noise level relative to the background noise alone was 1.65 dB.

b. The point estimate of signal-plus-noise level relative to the background during the peak of the envelope amplitude excursion was 3.25 dB.

c. The ratio of peak signal-plus-noise ratio to average signal-plus-noise ratio was, therefore, observed to be 1.6 dB.

The peak amplitude was measured using the transient capture mode of a real-time spectrum analyzer. This peak amplitude was then compared to the long-term average which included both the high amplitude and low amplitude signals.

After normalizing the average intensity to unity, the findings above show that the ratio of intensities at the peak of the amplitude to the average is:

$$(I + \Delta I)/I = 1.45/1.0$$

or

$$\Delta I/I = 0.45/1.0$$

$$10 \log (\Delta I/I) = -3.5 \text{ dB} .$$

Correspondingly, the confidence intervals are given by:

$$-4.3 < \Delta I/I < -3.3 \text{ dB} .$$

When this result is compared to those of Moore and Raab (1975), it is seen that the present observed SNR to respond agrees reasonably well. The pulses comprising the modulation fall between those investigated by them and, therefore, cannot be directly compared with the present results. However, extrapolating their results using the empirical method proposed by them, the predicted value of $\Delta I/I$ would be about -4.0 dB. The difference between this predicted value and the results obtained using the modified threshold procedure is certainly not unreasonable considering the simplifying assumptions. Hence, as for the classification of the high pass filtered signal, the results observed can be explained reasonably well by simply assuming that the problem is one of detecting a dichotomous feature. It must be noted, however, that treating these pulse-like increases in the amplitude as isolated pulses to be detected in a noise background

must be done with caution. These pulses occur often enough so that they are approaching an indistinguishable series of pulses as investigated by Miller and Taylor (1948). They state that the critical frequency is about 20 Hz and is also a function of duty factor.

In the absence of amplitude modulation, the listener is faced with the task of first detecting the presence of the unmodulated band of noise against the background. This detection is analogous to the case treated in Chapter VI where now W_{eff} is about 6000 Hz (Experiments VIII, IX, and X). Given the point estimate on $\overline{SNR}$ shown in Table 12 and the 1/3 octave calibration data, the effective signal-to-noise ratio is

$$S/N = -4.8 \text{ dB} \quad ,$$

which yields a detectability of

$$10 \log '_{opt} = 10.7 \quad .$$

This detectability compares favorably to that measured for the experiments where bands of noise were dichotomous (See also Figure 16). In this case, however, the difference between the SNR to detect the unmodulated noise and the SNR to make the terminal decision does not appear to be as great as that needed in Experiment VIII, although there is sufficient uncertainty about the effective bandwidth that a statistical test is not warranted.

For Experiment VI, the peak amplitude relative the long term average was measured to be about 2.0 dB although the natural noisiness of the envelope makes this quantity a difficult one to measure. From Table 12, Treatments VI.1 and VI.2 the average SNR is 2.78 dB, hence, the average signal-plus-noise to noise ratio 4.63 dB and the peak signal-plus-noise to noise ratio is 6.03 dB. Therefore, the peak signal-plus-noise ratio is 1.40 and

$$(I + \Delta I)/I = 1.38/1.0$$

or

$$\Delta I/I = .38/1.0$$

$$10 \log (\Delta I/I) = -4.2 \text{ dB} \quad .$$

This result is again near that extrapolated from the earlier experimental evidence. For this signal, however, the effective duration of the peak in envelope amplitude is more variable than that for Experiment V, and there is a considerable uncertainty about the value of pulse duration to use. However, the results are encouraging. When the amplitude modulation is absent, the $\overline{SNR}$ to respond is, from Table 12, about 5.9 dB when data for all sonar operators is pooled. This result yields a detectability of

$$10 \log d'_{opt} = 13.8$$

which is considerably higher than was noted for other experiments dealing with the detection of bands of noise. This result may be due to some residual periodic change in the spectrum of the signal, after

eliminating the amplitude modulation, which is heard as a modulation of the frequency structure of the signal. However, a difference in the classification decision rules could also be active and indeed this does seem to be the case in the experiment discussed below.

In Experiment VII, the uncertainties associated with the marine sounds are eliminated. The sound pattern in this case consists of two octave bands of stationary noise where the lower frequency one is either amplitude modulated by a 10 Hz square wave or is not. For this case, feature Ω_{4AM} has an effective duration of 50 msec and the peak in spectral level is 3 dB higher than the average for the modulated band.

Pooling data in Table 12, the required $\overline{SNR}$ when the modulated feature is present is 1.92 dB. From the calibration data and this observed mean signal-to-noise ratio to respond, the band level at the response was 4.7 dB relative to the background. The observation yields the following measures:

$$10 \log [(\overline{S} + n)/n] = 6.0 \quad ,$$

where $\overline{S}$ is the average signal intensity, and

$$10 \log [(S_p + n/N)] = 8.4 \quad ,$$

where S_p is the peak value of signal intensity, hence,

$$(I + \Delta I)/I = 1.73/1.0$$

or

$$10 \log (\Delta I/I) = -1.37 \quad .$$

Hence, for Experiment VII, the observed intensity discrimination performance was -1.4 dB for 354 Hz bandwidth and duration 50 msec. This result also falls within the anticipated performance bounds determined from data measured by Moore and Raab (1975).

When the amplitude modulation feature is absent, the problem reduces to detecting the band of noise centered at 500 Hz and then deciding that the amplitude modulation is absent. Comparing Tables 11 and 12, it is seen that the decision in favor of H_0 when the dichotomous feature was amplitude modulation is made at a $\overline{SNR}$ which is some 6 to 8 dB higher than when the dichotomous feature was the low frequency band of noise itself. The difference between these two sets of data clearly demonstrates that the listener alters the decision rules in response to the task to be performed. The detectability of the unmodulated band of noise is physically the same for Treatments IX.1 and IX.2 as it is for Treatments VII.3 and VII.4 but the response level is significantly different. The difference seems to be due to the fact that the listener must obtain sufficient information to decide that the band of noise being attended to is both present and not amplitude modulated.

7.4 The Effect of Distortion of the Signal Waveform

To test the effect of limiting on signals, the tape recorded signal was passed through a non-linear circuit with the transfer characteristic shown in Figure 18. The input level was so adjusted that limiting began at the 1σ point of the time waveform. Such a transfer function, while modifying the detailed amplitude vs time

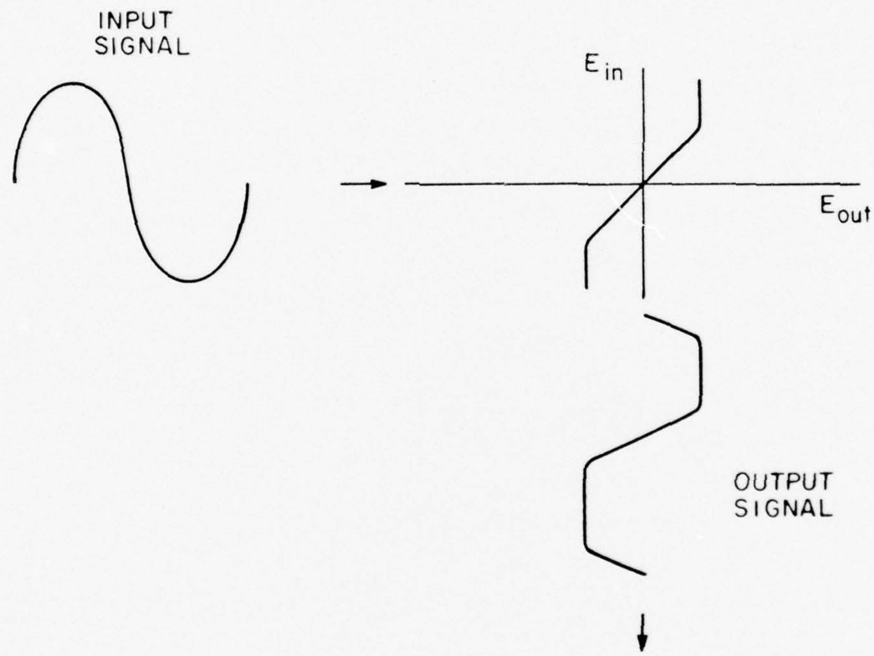


Figure 18. Time Domain Transfer Function of a Limiting Circuit Used to Vary the Temporal Structure of the Sound.

structure of the signal, does not greatly reduce the intelligibility of speech (Licklider, Bindra, and Pollack, 1948).

Five treatments using two marine source recordings were investigated using student subjects. Of the 53 events, only 18 resulted in a terminal decision and, of these, 55 percent were correct classifications. The treatments and number of events tested in each case are summarized in Table 14. The responses and listener comments indicate that there is little subjective difference in the sounds attributable to the limiting process. Those responses which were made occurred at a SNR near the maximum presented. At this high SNR, the signals are little contaminated by the background noise and the task is one of matching an essentially pure probe to the exposed signal set.

TABLE 14

SUMMARY OF TREATMENTS TO INVESTIGATE THE EFFECT OF 1σ LIMITING OF A MARINE SOUND

Treatment No.	Exposure Sequence A	B	Probe	No. of Events Tested
X1.1	ω_2	$\omega_{21\sigma}$*	ω_2	11
X1.2	ω_2	$\omega_{21\sigma}$	$\omega_{21\sigma}$	11
X11.1	ω_3	$\omega_{31\sigma}$	ω_3	10
X11.2	ω_3	$\omega_{31\sigma}$	$\omega_{31\sigma}$	11
X11.3	$\omega_{31\sigma}$	ω_3	$\omega_{31\sigma}$	10

*$\omega_{21\sigma}$ is the signal pattern limited at the 1σ point.

CHAPTER VIII

CRITERION AND RESPONSE BIASES

8.1 Order and Training Effects vs Criterion

The results presented in Chapters VI and VII demonstrated that the data are influenced by the experimental technique. In addition, that data hinted at a significant difference in the way sonar operators and trained University student listeners perform in the discrimination task.

The data is reviewed in this chapter in an effort to identify the effect various aspects of the experimental technique may have on the outcome of the experiment. The effect of time into the response period will be investigated. It is seen that one difficulty in using threshold schemes such as the "modified threshold procedure" is that it is difficult to separate the dependence of response SNR on the detectability of the dichotomous feature from time dependent shifts in criterion or from other factors which have a time dependence. The results from tests designed specifically to test the effect of time are presented in Section 8.2.

Then, in order to explicitly investigate the role of criterion on the outcome of discrimination tasks, the results of slide wire experiments are described in Section 8.3. These studies allow the mapping-out of performance as measured in terms of P(C) and P(W), the

probabilities of a correct or wrong decision, as the detectability changes.

8.2 Experimental Variable and Response Interaction

In earlier chapters, it was noted that there is an order effect which seems to be related to the sequence of exposure of the training signal set. The bias was observed in several of the experiments involving high pass filtering of the sound or other sounds for which the dichotomous feature is a low frequency band of noise, and indicated the following trends:

a. The order of exposure A, B followed by the probe A gave higher values of $\overline{SNR}$ than the order of exposure B, A followed by the probe A.

b. The order of exposure A, B followed by the probe B gave higher values of $\overline{SNR}$ than the order of exposure B, A followed by the probe B.

Note that the significant factor was not whether the probe immediately followed the corresponding exposure signal. The determining factor seemed to be in which order the exposure signals were presented. In these cases, A was the signal pattern containing the dichotomous feature, and B was the one with it missing.

The explanation for this effect is not at all clear since it is difficult to conceive of a memory organization for the data which would favor this type of behavior. This is certainly an effect which must be contended with in such experiments and one which should be

better understood. An understanding of the mechanism responsible would strongly influence subsequent experiments using the modified threshold procedure.

Another consideration in this type of experiment is that of the time factor. In studies of auditory memory span, it is routine to present an exposure signal followed by a shadowing stimulus prior to the probe presentation (Parkinson, Parks, and Kroll, 1971). The effect of the shadowing stimulus is quite pronounced when tests are done with isolated letters or phonemes in a non-rehearsal situation. The low SNR presentation at the beginning of the response period may be considered a shadowing stimulus and this may lead to an increase in SNR to respond as time to respond increases. The experimental design includes randomization of the starting SNR to reduce this effect as well as to avoid listeners becoming conditioned to respond after a certain time. However, the SNR and time are still strongly correlated. If the initial SNR were fixed, the correlation would be nearly exact since the SNR increases with time.

The design of the experiment was such that the starting SNR values could take on any one of five integer dB values and these were chosen at random using a uniformly distributed random number table. In the case where the SNR increases linearly at 1/4 dB per second, the correlation coefficient relating SNR to time is

$$\rho = \sigma_t / \sqrt{\sigma_t^2 + 26.67} \quad , \tag{8.1}$$

where σ_t is the standard deviation of the observed time to respond.[1] This correlation function was derived by noting that the starting SNR is independent of the time to respond. There are, therefore, two independent random variables, and the SNR is a function of both.

When the standard deviation on SNR to respond is S_{SNR} with a typical value of 2.5, the value of σ_t will be approximately 10. For this value, Equation (8.1) yields

$$\rho = \sigma_t / \sqrt{\sigma_t^2 + 26.67} = 0.89 \text{ when } \sigma_t = 10 \quad .$$

This is a high value of correlation and suggests the difficulty of separating SNR dependence from factors which effect t directly.

It was noted that sonar operators participating in the modified threshold procedure were noticeably disturbed whenever they failed to respond in an event. The observed number of non-response events was very small for this group of subjects as a result. It appears, therefore, that these subjects were assigning a high cost to deferring the decision once a certain time had elapsed during the response period. Prior to that time, they knew that the likelihood of their having time to respond was high. However, as time passed, they were more likely to be denied an opportunity to respond, and the cost for the deferral decision would, therefore, change in their view. Such a shift of criterion with time occurs in many real-world discrimination tasks as well and will confound the effort to measure performance at a fixed

---

[1] The correlation function was derived by Mr. D. W. Martin, a fellow graduate student assigned to support data reduction of related sonar operator data.

criterion such as is implied in the instruction "reasonably certain." Loeb and Binford (1970) investigated some of the dependencies with time of monitoring tasks although their time scales were considerably longer and there may not be much relationship to the present study. It is reasonable, however, to conjecture that some shift in criterion will occur as the response period progresses.

The large correlation between response time and SNR will result in a downward shift in $\overline{\text{SNR}}$ due to the tendency on the part of the listeners to lower their time to respond. There seems to be no totally acceptable way to eliminate the shift in criterion with time since making the response time itself longer is also undesirable because the listeners become restless during the period of time during which the signal is blanked. Even making the response times much longer does not insure that subjects will not shift their criterion as is shown by Loeb and Binford (1970) for listening sessions lasting hours.

To investigate the effect of time and shadowing in greater detail, an additional experiment (Experiment XIII) was performed using the modified threshold procedure. For this test, an additional 30-second delay was introduced between the start of the response period and the onset of ramping of the SNR. Since this period corresponded to an essentially noise-only situation, the stimulus during this time served as a masking one. Events for which this additional delay was present were mixed at random with those where the ramp started immediately and the subjects were not aware of which events would have this

characteristic. Only sonar operator subjects were used. Unfortunately, these added events corresponded to Experiment IX for which the data exhibited a pronounced dependence on exposure order. When data comparable to Treatment IX.4 was used as a comparison, however, the results indicated that the addition of the added delay reduces the SNR to make the terminal decision significantly. The T statistic was in this case

T = -3.19 for Experiment XI vs Experiment IX
Treatment 4 with 19 degrees of freedom.

This value of T is significant at the 0.01 level. Other treatments of Experiment IX could not be compared because of insufficient data. Also, pooling of data was not possible because of the exposure order effect.

This result dramatically demonstrates the effect of listener's shifting their criterion. Not knowing that this event would last longer than others, they chose to respond at a lower confidence rather than to possibly not respond at all. In retrospect, the design of this experiment pre-ordained the outcome. However, when these experiments were designed, there was no reason to suspect that listeners, especially trained sonar operators, would find it so difficult to maintain a criterion stressed repeatedly in the instructions to the subjects. Feedback in such tests would perhaps enforce a subject's criterion although Robinson and Watson (1972) report that for highly trained listeners, feedback might actually

be detrimental. Prior knowledge by the subject of which events have an additional shadowing period may also be beneficial since the subjects could then appropriately adjust their concept of how long they have till the response period may end.

Another experimental characteristic unique to sonar operator tests was that the exposure signals were the same for six consecutive events. Both the signal patterns and exposure order were held fixed and the exposure set was only presented briefly prior to another event once the set had been trained. The listeners, therefore, had six opportunities to refresh their memory of the signal patterns. In order to test for learning effects within such a group of events, SNR to respond was compared between the first event of a group and the last, between the first and fifth, etc. A statistically significant difference was found between the first event of a group and all others in Experiment III only. This finding was true for both low experience and high experience operators. It is possible that the marine sound in this case, having little amplitude modulation or other readily discernible characteristic, was very noise-like and was easily confused with the ocean ambient sound against which the signal pattern was presented in the response period. Lacking a recently reinforced mental image of the background sound, the first event resulted in false classifications at low values of SNR. Subsequent events, however, were tested in a situation where the refreshed signal patterns could be mentally compared to a rather recent exposure to the background. The subjects did, in fact, comment that they would have liked to be able

to recall the background noise at will for purposes of comparison. The pre-recorded nature of the signal sequences precluded such a refinement, however.

The above described intra-group effect does not explain the previously mentioned dependence on exposure order. It is conceivable that the ordering of events within a group could have favored certain combinations of exposures and probes at the beginning of the group and other combinations at the end. However, when the first event of a group is eliminated from consideration, the exposure order effect persists. Also, the exposure order effect occurred in most experiments with a band of noise dichotomous, whereas the intra-group effect was significant in only Experiment III.

The possible effect of operator fatigue was also investigated. Since sessions consisted of 18 to 24 events, events occurring at the end of a session may be biased by tiring of the subject. However, no statistically significant difference was found between like events associated with any experiment which occurred at the beginning of the session compared with those occurring at the end of a session. Similarly, no inter-group dependence could be established by application of statistical hypothesis tests on the SNR to respond.

Based on the above discussion, the effects induced by the experimental design are:

a. There is a consistent and significant effect of order of exposure for tests involving dichotomous bands of noise, but seemingly not for other signal patterns.

b. The SNR to respond is strongly correlated to response time and the latter is influenced by an apparently strong aversion to letting the event terminate without responding.

c. Uncertainty about the sound of the background noise in which the probe is presented can cause the first event in grouped data to give anomalous results. This problem is apparently confounded by the fact that two similar but not identical background noises were used in these tests.

d. The experimental results do not seem to depend on the position of an event within a session.

e. The effects of a masking stimulus or of short term memory decay were not measurable by the technique used. However, revisions of the method may allow for measuring these effects.

8.3 Results Using Continuous Rating Methods

In order to study the effect of criterion in more detail, three experiments using the slide-wire technique were conducted using Navy personnel only. Each experiment consisted of six events of grouped data as previously described. Of the data thus collected, only the last set of data using the sound patterns ω_2 and ω_{2H30} will be discussed below. The other data was obtained after the subjects were initially interviewed and were still in the process of adjusting to the situation.

The signal patterns and order of the exposure set corresponds to Treatments II.5 and II.6 listed in Table 7. That is, the signal patterns were presented in the order ω_2, ω_{2H30} with ω_2 or ω_{2H30} serving as the probe with the ratios of occurrence of the two being 4 to 2, respectively. A total of 70 useful events were collected using both groups of operators, those with extensive experience and those with relatively little.

The data was output to a multi-channel strip chart along with a voltage proportional to the SNR set on the balanced mixer. As the listeners adjusted their classification and degree of certainty, a voltage between 0 and 5 volts resulted. In the neutral position, centered, this voltage was 2.5 volts and decreased with increasing Confidence A classifications and increased with increasing Confidence B classifications.

In these tests, the subjects were instructed to attempt to make a classification as soon as possible and to feel free to change their mind. See also Appendix A where the text of the instructions is shown. The subjects did in fact revise early opinions in a number of instances as more data became available. Prior to taking data, the listeners were instructed to indicate what, in their opinion, constituted a decision of "reasonable certainty" and this rating was near 8 or higher for all listeners. Parallax in setting the slider of the linear digit potentiometer resulted in differences of $\pm$ 1 on the rating scale at the center point depending on the subject. This parallax, and coarseness of the rating scale contributed a hysteresis in the output of about one digit.

In analyzing the data, the tentative decision was noted after every 2-dB change in balanced mixer setting. Furthermore, decisions were scored as below:

a. Confidence ratings between 2A and 2B were considered as no response.

b. Confidence ratings between 2 and 6 in either the A or B directions were noted as low confidence classification decision of A or B, respectively.

c. Confidence ratings exceeding 6 were treated as high confidence decisions of A or B.

Subsequent analyses of the results considered low confidence answers as those for which the listener indicated either a low confidence or a high one, i.e., for which the subject was at least minimally committed to an A or B classification. For a high confidence answer, the listener had to have indicated a confidence level of 6 or more. If low confidence answers are used, a greater number of both correct and incorrect classifications are expected than when only high confidence classifications are admitted.

Classifications were grouped into two criteria even though a continuous rating scale is used. This grouping is made because a number of listeners will probably not apply criteria in consistent enough a fashion to allow pooling of data which is resolved into more categories. Finer categories are possible for a single listener, but it is expected that different listeners assign various meanings to numeric ratings (Robinson and Watson, 1972).

The data obtained from slide-wire tests is a classification which is a function of the SNR and the criterion. When the listener's rating is less certain than the specified criterion, no matter what his tentative decision (A or B), he can be thought of deferring his decision as described in Chapter II. It is assumed that a listener in this case would also not commit to a terminal decision which he had previously been instructed to make only when a certain criterion was satisfied. The present subjects indicated that a rating of about 8 corresponds to the verbal instruction "reasonably certain" during the calibration phase of the session.

The observed probabilities of various event-response categories are related to the detectabilities of the signal patterns (SNR's) and to the criterion (Egan, 1975). These probabilities define the receiver operating characteristics (ROC's) in these classifications. Figure 19 shows the salient features of an ROC curve for normal-normal distributions. The isocriterion contours correspond to a listener responding to signal patterns at various detectabilities while maintaining the same decision criterion. Shifts of criteria at fixed detectability result in translation parallel to the chance diagonal from one isocriterion contour to another. In order to visualize the various effects, consider the case where the hypotheses H_1 and H_0 are equally likely. Then, whenever the criterion is:

Respond either $\hat{\omega}_{.}$ or $\hat{\omega}_{.H..}$ in response to the stimulus.

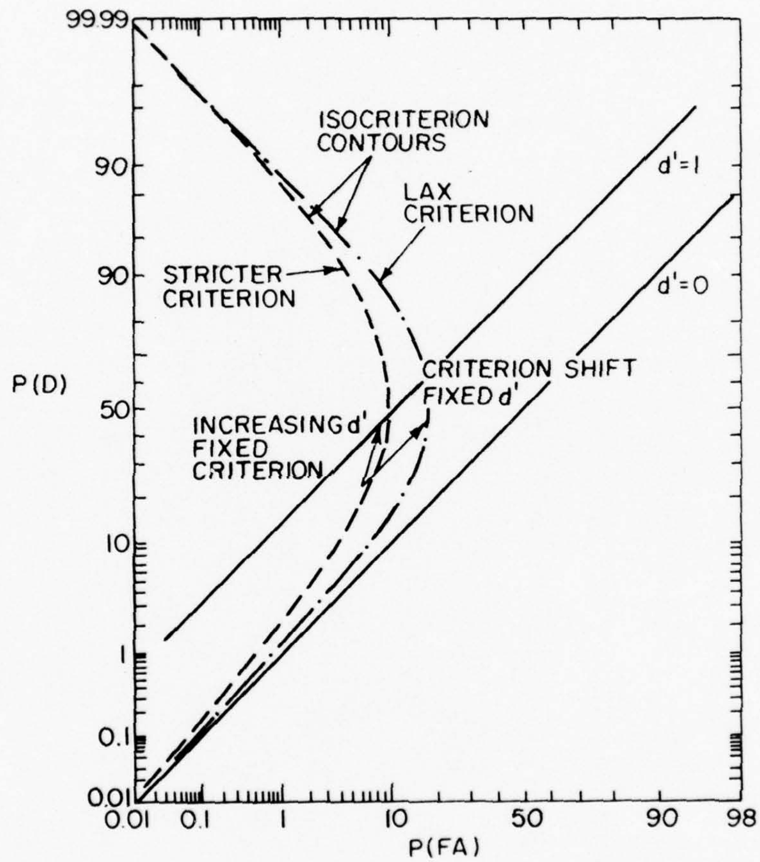


Figure 19. Salient Features of Receiver Operating Characteristic (ROC) Curves for Normal-Normal Distributions.

This is a lax criterion indeed since the response is made without necessarily any confidence. When the detectability of the signal is vanishingly small, the probability of correctly reporting the feature present, i.e., $P(\hat{\omega}_{.}|H_1)$ is equal to the probability of incorrectly reporting the feature present $(P(\hat{\omega}_{.}|H_0))$. Since there will be a response to each inquiry, the respective probabilities are each 0.5 since the feature is either present or not present with equal probability. In the figure, these probabilities are marked $P(D)$ and $P(FA)$ to conform to common practice. As the detectability increases, the relative probabilities change until there is no longer any uncertainty, and the presence of the feature is always detected and its absence never results in a positive report.

Similarly, when the criterion is:

Respond either $\hat{\omega}_{.}$ or $\hat{\omega}_{.H..}$ whenever reasonably certain,

both probabilities $P(D)$ and $P(FA)$ will be low but equal. As the detectability increases, a listener who wishes to maximize his performance in terms of $P(D)$ will shift the absolute value of the threshold to accept more false positive reports while increasing the relative probability of correct positive reports (Egan, 1975). As the detectability increases further, the value of $P(FA)$ will start to decrease as $P(D)$ continues to increase.

In the discrimination task, the performance indices are $P(C)$, the probability of a correct response, and $P(W)$, the probability of an incorrect response.

The probability of a correct response is given by

$$P(C) = P(\hat{\omega}_2|H_1)\ P(H_1) + P(\hat{\omega}_{2H30}|H_0)\ P(H_0) \quad , \tag{8.2}$$

where $P(\hat{\omega}_{.}|H_1)$ is the probability of a correct acceptance, $P(\hat{\omega}_{2H30}|H_0)$ the probability of correct rejection, and

$$P(H_1) = 2P(H_0), \quad P(\omega_2) = 2P(\omega_{2H30}) = .66$$

are the prior probabilities of the two signal patterns.

Similarly,

$$P(\omega) = P(\hat{\omega}_2|H_0)\ P(H_0) + P(\hat{\omega}_{2H30}|H_1)\ P(H_1) \quad . \tag{8.3}$$

Note that

$$P(C) + P(W) \leq 1.0$$

since these are in fact conditional probabilities conditioned on the criterion having been exceeded. The probabilities given by Equations (8.2) and (8.3) may serve as the variables when plotting the ROC curve and will exhibit the same dependence on d' and criterion as do $P(D)$ and $P(FA)$ in the detection problem.

Figures 20 and 21 show the experimental behavior of these probabilities with increasing SNR. Figure 20 is for the case where a strict criterion applies, and Figure 21 shows the results of accepting tentative decisions which satisfy the low confidence criterion or the high confidence one. The data includes the 90 percent confidence intervals on the estimate of underlying probability from the observed

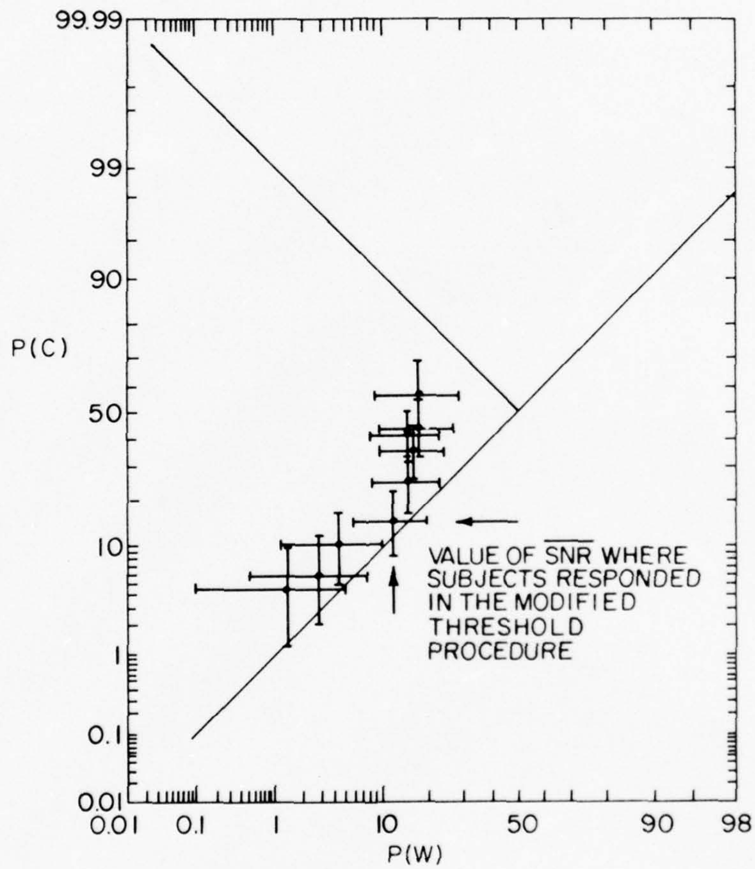


Figure 20. Experimental Results for Discrimination with a Strict Criterion for the Terminal Decision.

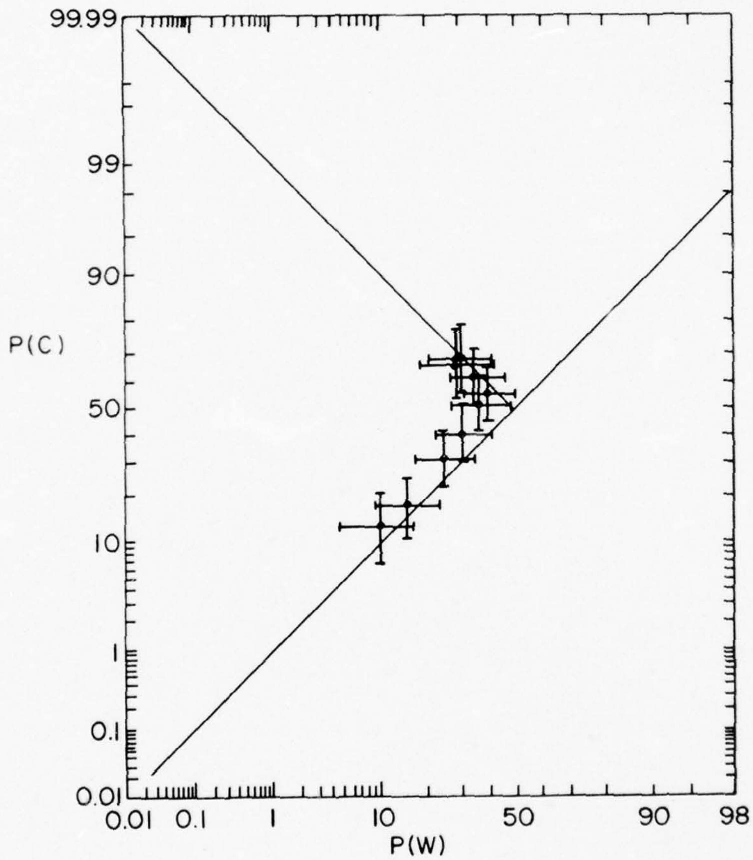


Figure 21. Experimental Results for Discrimination Under a Lax Criterion.

performance. Note that the two sets of data seem to follow the behavior predicted for a listener attempting to maximize his $P(C)$.

In order to more clearly show the effect of criterion, the results corresponding to an SNR of 1.0 and 11.4 are plotted in Figure 22 for the two criteria. The arrows show pure translation of the point estimates of $P(C)$ and $P(W)$ parallel to the chance diagonal as would be expected by a relaxation of the criterion. The observed behavior with these subjects is well within bounds at the 90 percent confidence level.

For the modified threshold procedure, the Navy personnel responded at an $\overline{SNR}$ of 6.46 when data shown in Table 7 is pooled. The criterion in this case was "reasonably certain," and this corresponds to ratings in the slide-wire tests of about eight. Therefore, the data collected in the slide-wire experiment under the strict response is comparable to the modified threshold data collected at the terminal decision point. In Figure 21, the approximately equivalent SNR data is shown with an arrow. Note that this datum occurs at the point where the $P(C)$ begins to increase rapidly in relation to $P(W)$.

It is also noteworthy that the value of $P(C)$ is not much different from the value of $P(W)$ at this $\overline{SNR}$. In fact, under the condition that a subject responds at this $\overline{SNR}$, the expected value of $P(C)$ for the modified threshold procedure is

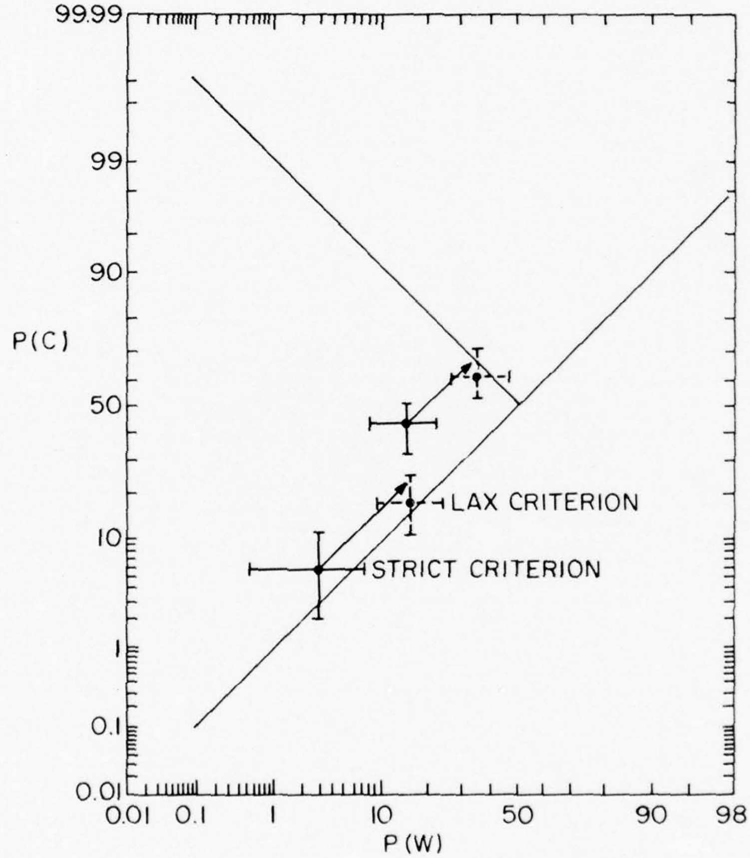


Figure 22. *Comparison of* Experimental Results for Strict and Lax Criteria Discrimination at Two Values of Signal-to-Noise Ratio.

$$P_{mt}(C|\text{Response}) = P_{SW}(C)/(P_{SW}(C) + P_{SW}(W))$$

$$= 0.243/(0.243 + 0.143)$$

$$= 0.63 \quad ,$$

which is comparable to values of $P(C)$ observed for other experiments using high pass filtering of marine sounds such as Experiment III. The subscripts mt and SW above are used to differentiate between the modified procedure and slide-wire experiment.

CHAPTER IX

SUMMARY AND CONCLUSIONS

The objective of the experiments reported in this thesis was to quantify the subjective response of listeners to noise-like sounds as are encountered frequently. Such sounds play a significant role in a number of situations ranging from the control of the speed of an automobile by the driver to tasking auditory discrimination tasks in industrial settings. The response of listeners to differences between sounds, the way in which they arrive at the classification decisions which are made routinely, and what the significant parameters of the sounds are, all are of both theoretical and practical interest. The importance of this information increases as there is a greater emphasis on protecting the worker from excessive noise through the use of hearing protectors and other means of isolating the human from such sounds. Also, the need to automatically monitor industrial processes places added emphasis on knowledge of how workers are able to interact with such processes by means of sounds.

The experiments consisted of eliciting discrimination decisions from listeners trained to recognize sound patterns. These patterns were both marine sounds and noise-like stimuli generated in the laboratory. Two experimental techniques were applied to populations of listeners consisting of trained University student subjects and of

Navy personnel trained in sound recognition. The experimental results are the required signal-to-noise ratio to discriminate between two sound patterns and the associated probabilities of correct or wrong classifications. In addition, a continuous rating scale was used to obtain more detailed information about the receiver operating characteristics (ROC's).

The treatise postulates a model of the discrimination task for a set of signal patterns which differ by either having or not containing a particular auditory feature, a dichotomous feature. In Chapter II, this model is presented in terms of a hypothesis test about the presence of the feature. It is suggested that, as a first approximation, the discrimination task can be treated as a detection problem where the significant factor is the detectability of the dichotomous feature. Known results from signal detection theory are used to obtain the optimum detection performance (d'_{opt}) for the case where the dichotomous feature is a band of noise at least as wide as the minimum critical auditory band.

In Chapter III, this model is expanded to treat the sequential recognition problem which is representative of many real-world auditory recognition tasks. In this situation, the listener will defer his decision until more information becomes available. The concept of a decision criterion is introduced and it is seen that this criterion plays a significant role both theoretically and, in later chapters, in experiments.

The experimental techniques used to investigate the listener performance is either a modification of classical threshold techniques, the "modified threshold" procedure, or a continuous rating method whereby the listener indicates his degree of confidence, the "slide-wire" experiment. The techniques and the various experiments performed are described in Chapters IV and V.

One of the principal areas of investigation was that of discrimination when the dichotomous feature is a band of noise. This is an area where a considerable amount of experimental data relating to the detection performance of humans is available. Also, the investigation is motivated by findings in speech recognition. Chapter VI details seven experiments which deal with this question. The dichotomous feature is either a band of noise resulting when one of the two signal patterns to be compared is high pass filtered or is a bandpass filtered noise when laboratory generated signals are used.

It is found that the model of the process as one of detecting the dichotomous feature gives good agreement with experimental results for some of the sound patterns used. The listener's decision to classify the sound when the dichotomous feature is present is in all cases made at a mean signal-to-noise ratio ($\overline{SNR}$) close to that predicted. The detectability is found to be

$$8.6 < 10 \log d'_{opt} < 11.3 \text{ dB}$$

with the probability of correct classification $P(C)$ averaging about 65 percent. This detectability is computed using the effective bandwidth of the dichotomous feature and an assumed integration time

of 500 msec. The predicted value of d'_{opt} is 9.2 dB at a P(C) of 0.65. When the feature is absent, the model must be revised to account for the fact that the listener must detect and act on other features of the sound which he uses as cues to decide that the dichotomous feature is absent.

In Chapter VII, the investigation of dichotomous features concentrates on temporal changes in the sound. The temporal, or envelope structure, is known to be important in speech perception. The experiments used sounds which differ in their amplitude modulation structure. For this dichotomous feature the agreement with the model of the process as one of detecting the feature is also found to be good. Where the required increment in intensity to the average $(\Delta I/I)$, the Weber fraction, is predicted to be -4 dB, the experimental results are in the range

$$-4.3 < \Delta I/I < -3.3 \text{ dB} .$$

In spite of the good agreement between a simple model of the process and the measured listener responses, the experiment identified some significant experimental biases and differences between populations of listeners. Chapter VIII deals with listener criterion and response biases. The criterion which the listeners use for making a final classification decision is strongly influenced by the experimental setting and their training. Significant interdependence is introduced by the modified threshold technique between the effect of signal detectability and that due to time into the response period.

The time seems to strongly impact the listener's criterion. The magnitude of this effect was not anticipated, especially for the population of listeners consisting of the Navy personnel, and some of the data is of marginal utility as a result. Unexpectedly, there is a significant difference in performance in the discrimination task between the student subjects and Navy personnel. The latter group of subjects consistently required a higher SNR to respond and were, in certain respects, not as accurate. A significant difference also exists between the Navy personnel with less than two years experience and those with extensive at-sea exposure.

The author set out to devise an experimental technique which could be used for the purpose of quantifying listener performance in a discrimination task and to apply the technique to a number of cases of practical interest in industrial situations. For this reason, the data is in some instances only adequate to do first order comparisons to the proposed model. However, good agreement is found between data collected by others in very specialized detection oriented situations and the present task which is a complex one in comparison.

A number of areas for further investigation have been identified. These include tests to resolve some of the ambiquities discussed in Chapter VIII. The relationship of continuous rating experiments to results obtained from the modified threshold technique is one area where a large volume of data is needed to decrease the uncertainties associated with the various probabilities associated with event-response alternatives. The effect of exposure order on the results

obtained for experiments with dichotomous bands of noise is an area which also warrants detailed investigation as it relates to the utility of such an experimental technique in studying the discrimination process. Although the modified threshold technique has been demonstrated to be a relatively quick way to obtain information about the terminal decision process, the strong interaction of time with the subject's criterion is a bothersome factor. A technique is needed which will allow time dependent criterion shifts to be identified separate from the effect of increasing signal detectability.

REFERENCES

Abramowitz, M., and I. A. Stegun, Handbook of Mathematical Functions, National Bureau of Standards, U.S. Department of Commerce, 1964.

Ahumada, A., R. Marken, and A. Sandusky, "Time and Frequency Analyses of Auditory Signal Detection," J. Acoust. Soc. Amer., 57:2, 1975, 385-390.

American Academy of Opthalmology and Otolarnygology, Committee on Conservation of Hearing, "Guide for the Conservation of Hearing in Noise," Los Angeles: Research Center, Subcommittee on Noise in Industry, 1957.

Annett, J., and L. Paterson, "Training for Auditory Detection," Acta Psychologica 27, 1967, 420-426.

Bauer, B. B., and E. L. Torick, "Researches in Loudness Measurements," IEEE Transactions on Audio and Electroacoustics, AU-14:3, 1966, 141-151.

Boring, E. G., A History of Experimental Psychology, New York: Appleton-Century-Crofts, 1950.

Corso, J. F., The Experimental Psychology of Sensory Behavior, New York: Holt, Rinehart and Winston, 1967.

Cremer, L., P. Plenge, and D. Schwarzl, "Kurven Gleicher Lautstarke Mit Oktav-Gefiltertem Rauschen Im Diffusen Schallfeld," Acustica 9, 1959, 65-75.

Davis, H., and S. R. Silverman (Ed.), Hearing and Deafness, New York: Holt, Rinehart and Winston, 1961.

Dubrovskii, N. A., and I. N. Tumarkina, "Investigation of the Human Perception of Amplitude-Modulated Noise," Soviet Physics, Acoustics, Vol. 13, No. 1, 1967, 41-47.

Durlach, N. I., "Note on Binaural Masking-Level Differences as a Function of the Interaural Correlation of the Masking Noise," J. Acoust. Soc. Amer., 1964, 1613-1617.

Egan, J. P., Signal Detection Theory and ROC Analysis, New York: Academic Press, 1975.

Environmental Protection Agency (EPA), "Laws and Regulatory Schemes for Noise Abatement," GPO Stock No. 5500-0046, 1972.

Eriksen, C. W., and H. W. Hake, "Multidimensional Stimulus Differences and Accuracy of Discrimination," J. Exp. Psychol., 50, 1955, 153-160.

Fletcher, H., "Auditory Patterns," Rev. Mod. Phys., 12, 1940, 47-65.

Fletcher, H., and W. A. Munson, "Relation Between Loudness and Masking," J. Acoust. Soc. Amer. 9, 1937, 1-10.

Fu, K. S., Sequential Methods in Pattern Recognition and Machine Learning, New York: Academic Press, 1968.

Goldstein, A. J., and J. R. Mundie, "Response of the Primary Auditory Neuron to Human Speech," Aerospace Medical Research Laboratory Report AMRL-TR-71-73, 1971.

Green, D. M., "Detection of Multiple Component Signals in Noise," J. Acoust. Soc. Amer., 30, 1958, 904-911.

a. Green, D. M., "Auditory Detection of a Noise Signal," J. Acoust. Soc. Amer., 32, 1960, 121-131.

b. Green, D. M., "Psychoacoustics and Detection Theory," J. Acoust. Soc. Amer., 32, 1960, 1189-1203.

Green, D. M., "Application of Detection Theory in Psychophysics," Proc. of the IEEE, 58, 1970, 713-723.

Green, D. M., and J. A. Swets, Signal Detection Theory and Psychophysics, New York: John Wiley and Sons, 1966.

Greenwood, D. D., "Auditory Masking and the Critical Band," J. Acoust. Soc. Amer., 33, 1961, 484-502.

Gundy, R. F., "Auditory Detection of an Unspecified Signal," J. Acoust. Soc. Amer., 33, 1961, 1008-1012.

Harris, J. D., Harris on Audiology, Groton, Conn., The Journal of Auditory Research, 1973.

Hawkins, J. E., Jr., and S. S. Stevens, "The Masking of Pure Tones and of Speech by White Noise," J. Acoust. Soc. Amer., 22, 1950, 6-13.

Helstrom, C. W., Statistical Theory of Signal Detection, New York: Pergamon Press, 1960.

Hogg, R. V., and A. T. Craig, Introduction to Mathematical Statistics, London: MacMillan, 1970.

Julesz, B., "Problems of Pattern Recognition and Perceptual Psychology," in L. N. Kanal (Ed.) Pattern Recognition, Washington, DC: Thompson Book Co., 1968.

Kaplan, H. M., Anatomy and Physiology of Speech, New York: McGraw-Hill, 1960.

Kinsler, L. E., and A. R. Frey, Fundamentals of Acoustics, New York: John Wiley and Sons, 1962.

Klatzky, R. L., Human Memory, Structures and Processes, San Francisco: W. H. Freeman, 1975.

Lathi, B. P., Random Signals and Communication Theory, Scranton, PA: International Textbook Co., 1968.

Leeper, R., "Cognitive Processes," in S. S. Stevens (Ed.), Handbook of Experimental Psychology, New York: John Wiley and Sons, 1951.

Lesser, V. R., R. Fennell, L. Erman, and D. R. Reddy, "Organization of the Hearsay II Speech Understanding System," IEEE Transactions on Acoustics, Speech and Signal Processing, ASSP-23:1, 1975, 11-23.

Licklider, J. C., "Basic Correlates of the Auditory Stimulus," in S. S. Stevens (Ed.), Handbook of Experimental Psychology, New York: John Wiley and Sons, 1951.

Licklider, J. C. R., D. Bindra, and I. Pollack, "The Intelligibility of Rectangular Speech-Waves," Amer. J. Psychol., 61, 1948, 1-20.

Loeb, M., and J. R. Binford, "Examination of Some Factors Influencing Performance on an Auditory Monitoring Task with One Signal Per Session," J. Exp. Psych., 83:1, 1970, 40-44.

McGill, W. J., "Polynomial Psychometric Functions in Audition," J. Mathematical Psych., 5, 1968, 369-376.

Miller, G. A., "The Perception of Short Bursts of Noise," J. Acoust. Soc. Amer., 20:2, 1948, 160-170.

Miller, G. A., and W. G. Taylor, "The Perception of Repeated Bursts of Noise," J. Acoust. Soc. Amer., 20:2, 1948, 171-182.

Moore, B. C. J., and D. H. Raab, "Intensity Discrimination for Noise Bursts," J. Acoust. Soc. Amer., 57:2, 1975, 401-406.

Murdock, B. B., and J. C. Ogilvie, "Binomial Variability in Short-Term Memory," Psych. Bull., 70:4, 1968, 256-260.

Murrell, K. F. H., Human Performance in Industry, New York: Reinholt, 1965.

Myasnikov, L. L., and Ye. N. Myasnikova, Automatic Recognition of Sound Images, Translated from Russian by the National Technical Information Service, Leningrad: Energiya, 1970.

Olach, J. D., "Visual Memory Span of Young Children for Sequentially Presented Meaningful Material," Pennsylvania State University Masters Thesis, September 1964.

Parkinson, S. R., T. E. Parks, and N. E. A. Kroll, "Visual and Auditory Short-Term Memory," J. Exp. Psych., 87:2, 1971, 274-280.

Perrone, A. J., "Ambient-Noise-Spectrum Levels as a Function of Water Depth," J. Acoust. Soc. Amer., 48:1, 1970, 362-370.

Peterson, A. P. G., and E. E. Gross, Handbook of Noise Measurement, West Concord, Mass.: General Radio Co., 1963.

Peterson, W. W., T. G. Birdsall, and W. C. Fox, "The Theory of Signal Detectability," Trans. I.R.E., PGIT-4, 1954, 171-212.

Plomp, R., and G. F. Smoorenburg, (Ed.), Frequency Analysis and Periodicity Detection in Hearing, Leiden: A. W. Sijthoff, 1970.

Pollack, I., "Effects of High Pass and Low Pass Filtering on the Intelligibility of Speech in Noise," J. Acoust. Soc. Amer., 20:3, 1948, 259-266.

Pollack, I., "Identification of Elementary Auditory Displays and the Method of Recognition Memory," J. Acoust. Soc. Amer., 31, 1959, 1126-1128.

Raab, D. H., and I. A. Goldberg, "Auditory Intensity Discrimination with Bursts of Reproducible Noise," J. Acoust. Soc. Amer., 57:2, 1975, 437-447.

Raisbeck, G., Information Theory, An Introduction for Scientists and Engineers, Cambridge, Mass.: MIT Press, 1963.

Reed, S. K., Psychological Processes in Pattern Recognition, New York: Academic Press, 1973.

Robinson, D. E., and C. S. Watson, "Psychophysical Methods in Modern Psychoacoustics," in J. V. Tobias (Ed.), Foundations of Modern Auditory Theory, Vol. II, New York: Academic Press, 1972.

Shannon, C. E., and W. Weaver, The Mathematical Theory of Communication, Urbana, Ill.: University of Illinois Press, 1949.

Sheridan, T. B., and W. R. Ferrell, Man-Machine Systems: Information, Control, and Decision Models of Human Performance, Cambridge, Mass: MIT Press, 1974.

Sholl, H. A., "Man-Machine Systems for Detection, Recognition, Transmission, and Perception of Information," Vol. II, Part III, Connecticut University Report, 1971.

Siebert, W. M., "Frequency Discrimination in the Auditory System: Place or Periodicity Mechanisms," Proc. of the IEEE, 58, 1970, 723-730.

Sorkin, R. D., "Extension of the Theory of Signal Detectability to Matching Procedures in Psychoacoustics," J. Acoust. Soc. Amer., 34, 1962, 1745-1751.

Stallard, J. M., and C. B. Leslie, "Psychoacoustics and Passive Sonar Detection," Naval Ordnance Laboratory Technical Report, NOLTR 74-27, 1974.

Stevens, K. N., and A. S. House, "Speech Perception," in J. V. Tobias (Ed.), Foundations of Modern Auditory Theory, Vol. II, New York: Academic Press, 1972.

Stevens, S. S., "Neural Events and Psychophysical Law," Science 170, 1970, 1043-1050.

Swets, J. A. (Ed.), Signal Detection and Recognition by Human Observers, New York: John Wiley and Sons, 1964.

Swets, J. A., and D. M. Green, "Sequential Observations by Human Observers of Signals in Noise," in C. Cherry, (Ed.), Information Theory, London: Butterworths, 1961.

Swets, J. A., D. M. Green, and W. P. Tanner, Jr., "On the Width of Critical Bands," J. Acoust. Soc. Amer., 34, 1962, 108-113.

Swets, J. A., and S. T. Sewall, "Stimulus vs. Response Uncertainty in Recognition," J. Acoust. Soc. Amer., 32, 1961, 1586-1592.

Swets, J. A., E. F. Shipley, M. J. Mikey, and D. M. Green, "Multiple Observations of Signals in Noise," J. Acoust. Soc. Amer., 31, 1959, 514-521.

Swets, J. A., W. P. Tanner, Jr., and T. G. Birdsall, "Decision Processes in Perception," Psychol. Rev. 68, 1961, 301-309.

Tanner, W. P., Jr., "Application of the Theory of Signal Detectability to Amplitude Discrimination," J. Acoust. Soc. Amer., 33, 1961, 1233-1244.

Tanner, W. P., Jr., "Theory of Signal Detectability as an Interpretive Tool of Psychophysical Data," J. Acoust. Soc. Amer., 32, 1960, 1140-1147.

Tanner, W. P., and R. D. Sorkin, "The Theory of Signal Detectability," in J. V. Tobias (Ed.), Foundations of Modern Auditory Theory, Vol. II, New York: Academic Press, 1972.

Treisman, M., and F. R. Watts, "Relation Between Signal Detectability Theory and the Traditional Procedures for Measuring Sensory Thresholds," Psychol. Bull., 66:6, 1966, 438-454.

Van Trees, H. L., Detection, Estimation, and Modulation Theory, Part I, New York: John Wiley and Sons, 1968.

Wald, A., Sequential Analysis, New York: John Wiley and Sons, 1947.

Walpole, R. E., and R. H. Myers, Probability and Statistics for Engineers and Scientists, New York: MacMillan, 1972.

Ward, W. D., and J. E. Fricke, (Eds.), "Noise as a Public Health Hazard," OSHA Report No. 4, 1969.

Weinberg, G. H., and J. A. Schumaker, Statistics an Intuitive Approach, Benton, CA: Wadsworth Publishing Co., 1962.

Winer, J. B., Statistical Principles in Experimental Design, New York: McGraw-Hill, 1962.

Zhukov, S., and L. A. Christovitch, "Some New Concepts in Auditory Analysis of Acoustic Flow," Sov. Phys. Acoust., 20:3, 1974, 237-240.

Zwicker, E., "Ueber Psychologische und Methodische Grundlagen der Lautheit," Acustica, 8:1, 1958, 237-258.

Zwicker, E., G. Flottorp, and S. S. Stevens, "Critical Bandwidth in Loudness Summation," J. Acoust. Soc. Amer., 29, 1957, 548-557.

Zwislocki, J., "Temporal Summation of Loudness: An Analysis," J. Acoust. Soc. Amer., 46, 1969, 431-441.

APPENDIX A

DOCUMENTS AND INSTRUCTIONS TO SUBJECTS

QUESTIONNAIRE - FORM 1

A STUDY OF RADIATED NOISE (________________)

Note: This material will be treated as personal, please answer all questions fully.

Name: ______________________________ Grade: ________________________

Current Billet: ____

Approximate number of years of sonar experience ____________________

Does this experience include operations involving surface ship contacts? ____

List recent shore based sonar schools attended

Year	Course	Where
____________	____________________________	____________
____________	____________________________	____________
____________	____________________________	____________
____________	____________________________	____________

* * * * *

Medical History: Do you now have: Ringing in the ears ______, A head cold ______, perforated ear-drum ____________.

Have you had in the past the following symptoms?

Symptom (x)	Often	When you were young	Recently
Ear aches	( )	( )	( )
Discharge from the ear	( )	( )	( )
Dizziness	( )	( )	( )
Trauma due to explosion	( )	( )	( )
Temporary loss of hearing following exposure to gun-fire or other loud noise	( )	( )	( )

Questionnaire Used with Navy Personnel.

INFORMED CONSENT AGREEMENT -
PARTICIPATION IN A STUDY OF RADIATED NOISE

You will be participating in an experiment designed to test the ability of a listener to respond to changes of sounds. The test will be conducted in a quiet enclosure using headphones, and the loudness with which the sounds will be presented will be carefully controlled. At no time will the sound be so loud as to cause discomfort but if, at any time, you feel uncomfortable you should remove the headphones and leave the test enclosure. Your doing this will not be used as a basis for discontinuing your participation in the test. You may also terminate your participation in the experiment at any time.

For your convenience, test sessions may be scheduled at most reasonable times. However, to avoid fatigue and to keep from biasing the experiment, no more than one 60 minute test session may be done in any 24 hour period.

The sounds to which you will be listening may or may not be familiar to you. The method for recording your response, the order in which signals are presented, and the details of the experiment will be explained to you prior to beginning any test session. If you have any questions, please ask for clarification by the experimentor. A memorandum, "Prospectus for a Study of Radiated Noise as a Communicative Signal," is available for you to read if you desire to know more about the overall test objectives, its application and rationale.

I, ______________________________, have read this document
subject signature

to understanding. Date ______________________

Witnessed by: ______________ Date ______________________

Informed Consent Agreement for University Student Subjects.

Instructions appearing on the first audio tape to which Navy personnel were exposed. These instructions were followed by a question/answer period during which time all uncertainties about the experimental technique were to be resolved.

> You will be participating in an experiment designed to test the ability of a listener to detect changes in sound. The loudness at which the sounds are presented will be carefully controlled. At no time will the sound be so loud as to cause discomfort, but if at any time you feel uncomfortable you should remove the headphones and make this fact known to the test conductor. Also, if you are at all uncertain as to the test procedures you should ask for clarification.
>
> The sounds you will hear will be marine noises which may or may not be familiar to you. You will be asked to learn these sounds and then decide which is presented in the noise. The test sequence will consist of two signals presented without interfering noise. These signals will be denoted signal A and signal B. Signal A will be presented then signal B. The signal will then be repeated A - B. You will then be presented signal A or signal B in a noise. This is a response period which will be indicated by a green light on the response recorder. The objective is to indicate your decision of whether the unknown signal is signal A or signal B. The criterion you should use in making your decision, as well as the way your response is to be indicated, will vary. You will always be instructed at the beginning of a test session about the detailed procedure to be followed. For this series of experiments, the tests are organized into groups of events. For all the events of a group, the signals A and B will be the same. After first being presented in the order A then B, and repeated in the sequence A - B, they are only repeated once more in this order to refresh your memory of the sounds before another response period begins. The sounds A and B are presented randomly in the noise with the likelihood of signal A being presented in the noise being equal to the likelihood of signal B being presented in the noise.
>
> An example of a test sequence follows. In this case, the amount of noise mixed with the signal will decrease slowly. You will notice that the sound is at first masked by the noise. As the amount of noise changes, it will become easier to hear the sound. Finally, you will be able to decide with reasonably certainty which sound, A or B, it was which was presented with the noise. Listen to this sequence and try to decide which sound is presented in the noise.

Instructions on the audio tape to which Navy personnel were exposed prior to beginning tests using the slide-wire continuous rating response recorders.

For this experiment the response recorder will be used to indicate your degree of certainty about which sound it is which is presented in the noise. Initially, the indicator should be centered on zero. The amount of noise mixed with the signal will decrease slowly. Move the indicator to record your decision about which signal is the unknown. If you think perhaps that the unknown signal is signal A, move the indicator up the scale towards A. The further up the scale, the more certain you are of your choice. Try to make a tentative decision as soon as you can and do not hesitate to change your mind as more information becomes available.

You should try to associate a numeric value with your degree of certainty. For example, an 8 should be used to indicate a reasonably certain decision. For calibration, now indicate a response as if you were reasonably certain that the unknown signal was B. Now indicate a low confidence of a decision of A, and a high confidence, very certain, of a decision of the unknown being A. At the conclusion of an event, always return the indicator to the zero position. Reset the indicator to zero now. Please indicate your responses for the following events. Remember, that for all events in a group the signals A and B are the same. Also, try to maintain the same numeric value for your criteria and recenter the indicator between events.

Instructions recorded at the beginning of all auditory tapes used with Navy personnel. These tests all used grouped data.

> The test sequence will consist of two signals presented without interfering noise. These signals will be denoted signal A and signal B. Signal A will be presented, then signal B. The signals will then be repeated in the sequence A then B. During the response period, which will be indicated by a green light on the response recorder, either signal A or signal B will be presented in a noise. The amount of noise will decrease slowly. The objective is to indicate your decision as to which signal you conclude is mixed with the noise. Indicate your choice by pressing the switch marked A if you decide that signal A was mixed with the noise, or, press the switch marked B if you decide that signal B was mixed with the noise. You should indicate this decision as soon as you can under the condition that you are reasonably certain of your choice. For this series of experiments the tests are organized into groups of events. For all events of a group the signals A and B will be the same. The signals A or B are presented randomly in the noise with each being equally likely. Now, please indicate your classification decision for the following cases.

Voice comments following grouped events and leading into another group or terminating the test session.

This is the end of one group of tests. Another group of tests follows. For this group the signals A and B will be the same. However, these will generally not be the same signals as in the previous group. Please try to learn these signals without regard for other signals you have heard during this experiment.

This is the end of another group of tests. In the group of events which follows, the signals A and B are the same throughout. These signals will generally not be the same as those heard previously. Please try to learn these sounds independent of other signals you may have been exposed to in this experiment.

End of another group of events. The final group of events follows. For this group as in those before, the signals A and B will be the same. The likelihood of signal A being presented in the noise is the same as is the likelihood of signal B being presented in the noise.

This concludes a test session. Thank you for your cooperation.

APPENDIX B

HARDWARE SYSTEM

The tapes used as the source of the auditory stimuli in these tests are the end product of a carefully controlled series of taping steps. The marine sources of interest come recorded on a large variety of machines, at various tape speeds and recording techniques. Selected portions of these raw source tapes are re-recorded onto 1/2 inch magnetic tapes in FM recording mode and at 60 or 30 inches per second. During the re-recording of these master tapes, adjustment is made for hydrophone or other frequency weighting appearing on the raw data. The signals are also prewhitened by an amplifier with a 6 dB per octave increasing gain vs frequency behavior.

The master tapes serve as inputs to any stimuli recordings which use actual recorded marine sounds. Alternately, a weighted noise source may be substituted for either member of the exposure set. Recorded ocean ambient or shaped noise serves as the background noise against which the probe signal is presented. These various signals are properly time multiplexed by means of an analog selector which receives control signals from a digital sequencer. Figure 23 shows the various components of the system needed to create primary tapes from master tapes. The flow of analog and digital signals is also shown. In Figure 23, those components which are not specifically

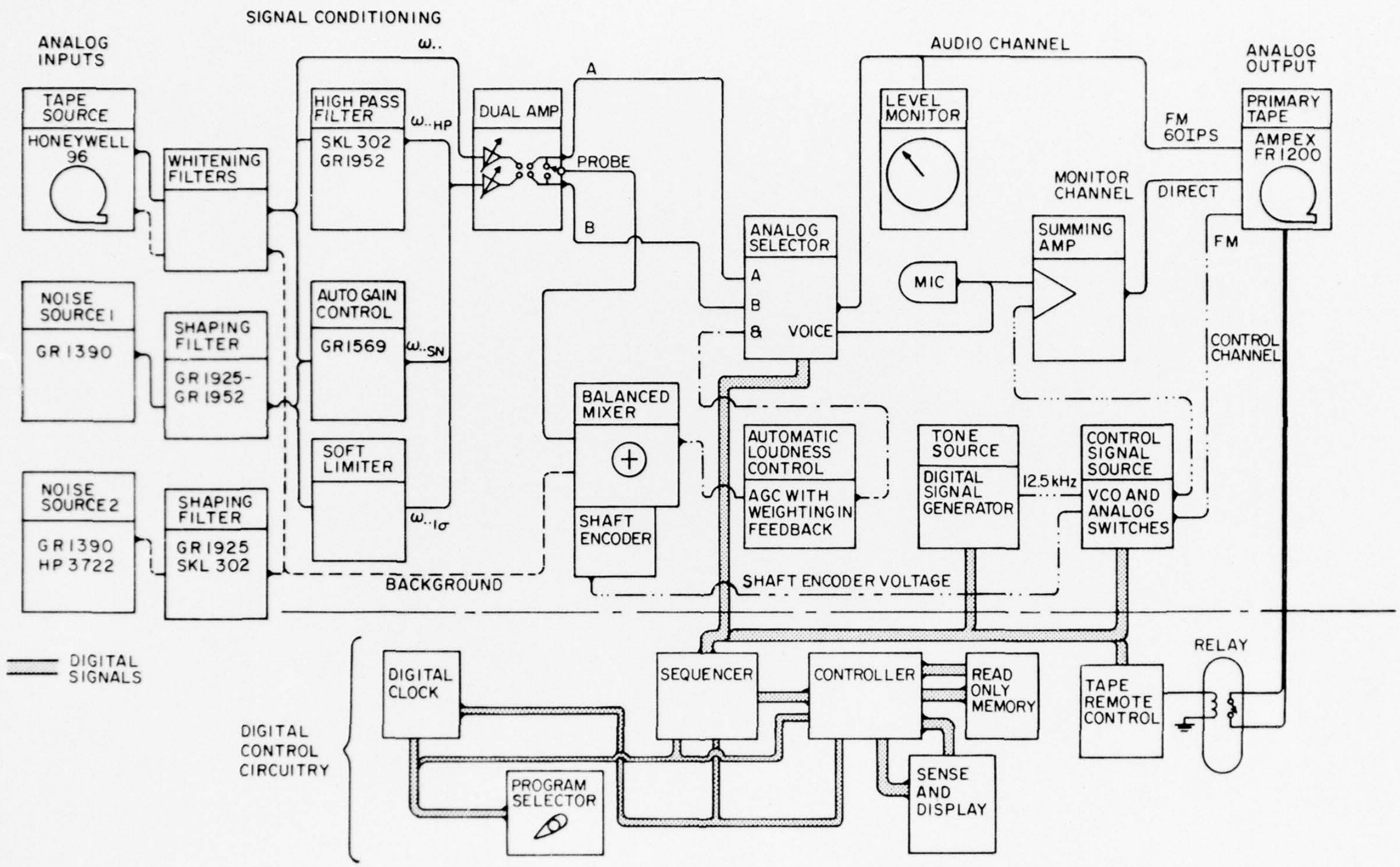


Figure 23. Flow Diagram Showing the Principal Components Needed to Create Primary Tapes.

indicated as being commercial items were fabricated at the Applied Research Laboratory. Most of these components were designed and built by the author. The major portions of the system with the exception of the analog input tape unit are pictured in Figure 24. Necessary interconnecting cabling has been omitted from this photograph, however.

The digital sequencer and associated controller orchestrates the required functions in response to a simple program stored in a read only memory. The sequencer starts and stops the output tape drive, sequences the various audio signals, and causes the generation of control signals which will be used at the test site to indicate response periods and to record responses. The exact order of the process is determined by a program selector and by the state of sense switches. The setting of the balanced mixer is performed manually in response to a cue light which indicates the response period portion of the audio sequence.

A direct recorded monitor channel of the output tape is used for voice annotation of cut numbers on the primary tapes and to record a 12.4 kHz pilot tone. This tone is on during valid portions of data on an FM recorded audio channel and an FM recorded control channel. This control channel has tones for control of the response recorder and a balanced mixer setting proportional VCO output.

Cuts from the primary tapes are re-recorded onto the final audio tapes in a randomized order. The monitor channel of the primary tape is used both to locate the desired cuts and to allow for voice

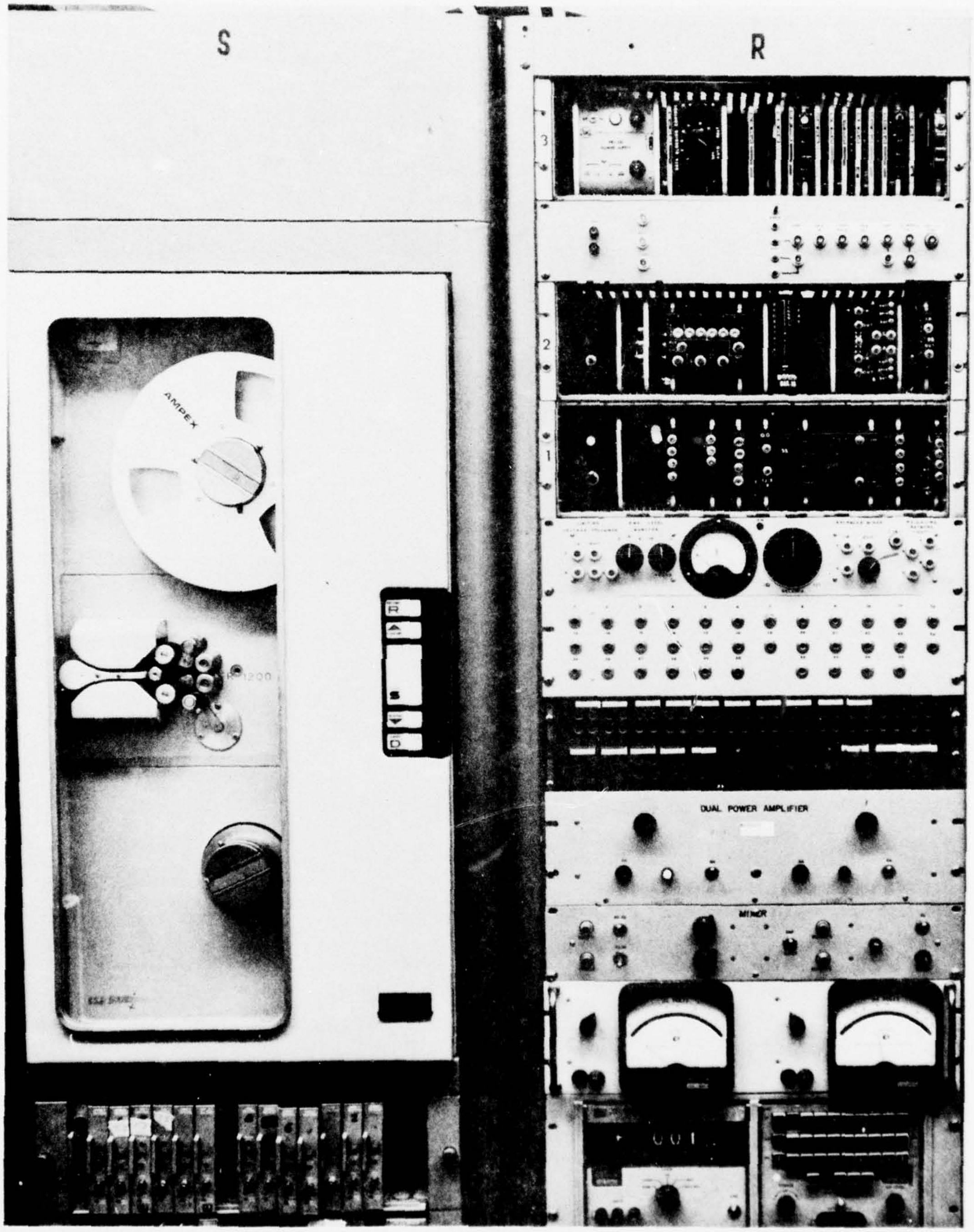


Figure 24. Photograph of the Hardware System Used to Create Primary Tapes with Interconnecting Cabling Omitted.

annotation between events on the audio tapes. This voice annotation of event numbers and of special instructions to the listener is usually from a cassette recording. A phase locked loop, analog selector combination, insures that the objectionable FM discriminator noise output while searching the primary tape does not appear on the tapes used for listener tests. Figure 25 diagrams this aspect of the recording process.

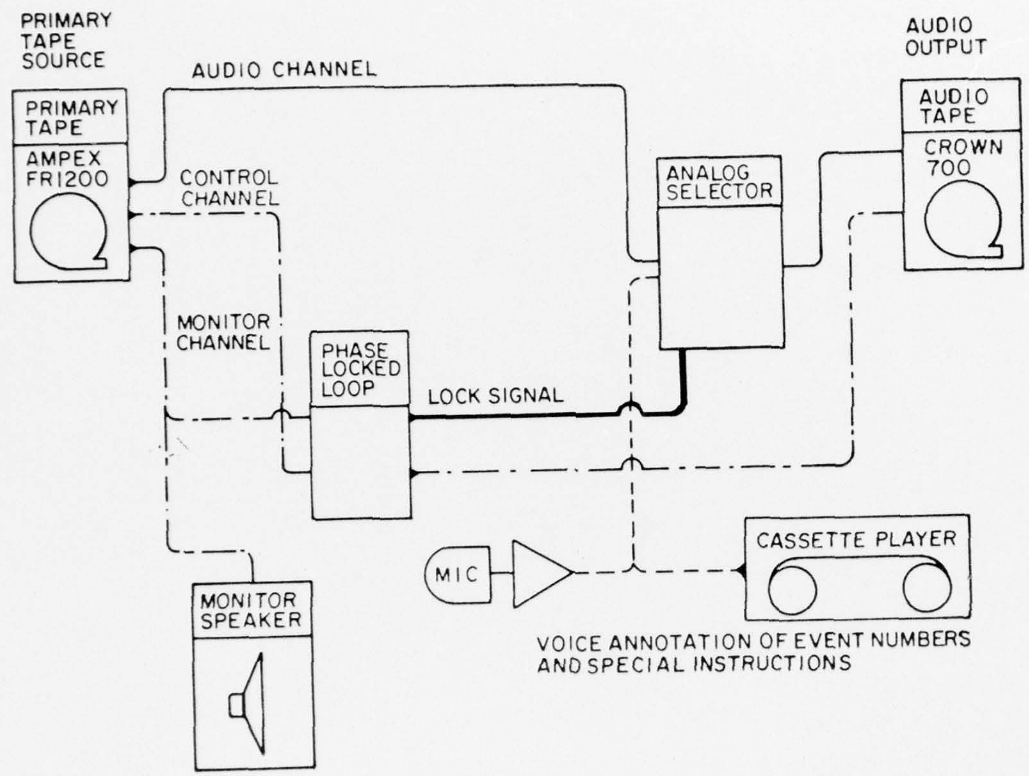


Figure 25. Schematic Diagram of the Components Used in the Recording of Audio Tapes in the Experiments in Psychoacoustics.

VITA

Claus Peter Janota was born Klaus P. Schneider in Munich, Germany, on 6 September 1941 and was granted U.S. Citizenship on 20 October 1955. He graduated from Wichita Falls, Texas, Senior High School in May 1959, and received the Bachelor of Science Degree in Physics and Mathematics from Midwestern University, Wichita Falls, Texas, in September 1962, as the first physics graduate from that institution.

Mr. Janota was a graduate assistant in the Field Emission Laboratory of The Pennsylvania State University until Summer 1964, at which time he entered the U.S. Army as an officer in the Ordnance Corps. He served in Germany at Wildflecken, and later as 7th Army Conventional Ammunition Supply Officer in Kaiserslautern. He was awarded the Army Commendation Medal for his role in Theatre Supply Automation.

In March 1968, he entered employment at the Ordnance Research Laboratory of The Pennsylvania State University working in the field of digital signal processing. In September 1971, he received the Master of Science Degree in Engineering Acoustics.

Since 1972 he has been working in systems and signal analysis under contract with the Naval Air Systems Command at the Applied Research Laboratory. Also, since 1973, he has served as a consultant to the Naval Sea Systems Command in Undersea Combat Systems.

He is a member of Sigma Xi.

DISTRIBUTION

Commander (NSEA 09G32) Naval Sea Systems Command Department of the Navy Washington, D. C. 20362	Copies 1 and 2
Commander (NSEA 0342) Naval Sea Systems Command Department of the Navy Washington, D. C. 20362	Copies 3 and 4
Defense Documentation Center 5010 Duke Street Cameron Station Alexandria, VA 22314	Copies 5 through 16